Universidad
Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Essay on Contests and Voting

Autor:

Marco Serena

Director/es:

Luis Corchón

DEPARTAMENTO DE ECONOMIA

Getafe, junio 2016

(a entregar en la Oficina de Posgrado, una vez nombrado el Tribunal evaluador , para preparar el documento para la defensa de la tesis)

TESIS DOCTORAL

ESSAY ON CONTESTS AND VOTING

Autor: ***Marco Serena***

Director: **Luis Corchón**

Firma del Tribunal Calificador:

Firma

Presidente: (Nombre y apellidos)

Vocal: (Nombre y apellidos)

Secretario: (Nombre y apellidos)

Calificación:

Leganés/Getafe, de de

Acknowledgements

I would not have been able to write this thesis without the constant presence of a number of people around me.

First of all, I would like to thank my adviser Luis Corchón for his tireless willingness to guide me through the process of doing research and for having taught me that being hardworking, focused and humble pays back. A key role was also played by Carmen Beviá, whose valuable insights significantly contributed to the process of writing this thesis. Furthermore, useful comments and feedback have been provided by several other economists, including Guillermo Caruana, Philipp Denter, Tore Ellingsen, Christian Ewerhart, Natalia Fabra, Qiang Fu, Aart Gerritsen, Andrea Guariso, Angel Hernando-Veciana, René Kirkegaard, Jingfeng Lu, Thomas Mariotti, Ricardo Martinez, Diego Moreno, Ron Siegel and Giancarlo Spagnolo.

Furthermore, I would like to thank my officemates, who made my routine much happier: in chronological order, Carlos Marín Gallardo, nobody, Salvatore Lo Bello, Lian Allub, Federico Masera, Mehdi H. S. and Cleiton Guollo. They taught me a lot. In particular, thanks to Salvatore I know that there are no arbitrage opportunities in online betting; Lian made me taste the stress of the job market, which was a useful experience; Federico patiently listened to the stories of when I played tennis against Nadal; Mehdi convinced me that econometricians know less than stock market agents.

Several friends constantly reminded me that there is much more outside academia to enjoy: Alex Barrachina, José Manuel Carbo, Federico Curci, Emre Ergemen, Andrés García-Suaza, Beatriz Gonzalez, Matthias Kredler, Iacopo Morchio, Sebastian Panthoefer, Anıl Yıldızparlak and Pia Zanetti. A special thanks to Pedro Sant'Anna for introducing me to the world of vertical steamers, to Christos Mavridis for teaching me how to work with binomial series, and to Fabrizio Colella for teaching me that sometimes it might be best to just close your eyes to problems.

Jan Bietenbeck, Alessandro Galesi, Alessandro Peri, Omar Rachedi and Francesco Risi have provided a steady and reliable source of fun, food and help, whenever needed. A burgeoning number of good memories of my stay in Madrid would not exist without them. I also thank Claudia Pochetti for years of unconditional support.

I owe my family nearly everything good I am and have.

Abstract

This thesis has three chapters. In the first chapter I study the role of information in contests. A contestant's effort depends on her knowledge of her rival's type. This knowledge is often limited in real-life contests. I propose a model where the principal of a contest has commitment power to verifiably disclose contestants' types. I investigate the optimal disclosure policy to stimulate contestants' efforts. I find that full disclosure spurs more (less) efforts than full concealment if the distribution of types is skewed toward high- (low-) types. However, the optimal disclosure policy is a particular partial disclosure, regardless of the skewness of the distribution of types; it consists of disclosing the signal which is best for the principal (i.e., all contestants are high-types) and concealing the rest.

In the second chapter I propose a novel objective function of a contest designer. When the winner selection process in a contest is noisy, the designer should take this noise into consideration when designing the contest if her goal is to maximize the quality of the winning entry. I propose an objective function that accommodates this idea, and I compare the optimal contest design under this objective function to the one under the commonly assumed maximization of sum of contestants' efforts. I find that, contrarily to what happened when the designer maximized the sum of efforts, the optimal contest design changes in that a designer may now benefit from: unlevelled playing field, exclusion of weak contestants, and weakening of the underdog.

In the third chapter, which is a joint work with Christos Mavridis, we contribute to the literature on pivotal voter models. For small electorates, the probability of casting the pivotal vote drives one's willingness to vote, however the existence of costs of voting incentivizes one's abstention. In two-alternative pivotal-voter models, this trade-off has been extensively studied under private information on the cost of voting. We complement the literature by providing an analysis under complete information, extending the analysis of Palfrey and Rosenthal [1983. A strategic calculus of voting. Public Choice. 41, 7-53]. If the cost of voting is sufficiently high at least for supporters of one of the two alternatives, the equilibrium is unique, and fully characterized. If instead the cost of voting is sufficiently low for everyone, we characterize three classes of equilibria and we find that all equilibria must belong to one of these three classes, regardless of the number of individuals. Furthermore we focus on equilibria which are continuous in the cost of voting. We show that this equilibrium refinement pins down a unique equilibrium. We conclude by discussing an application of our findings to redistribution of wealth.

Contents

Chapter 1
Harnessing Beliefs to Stimulate Efforts

1 Introduction

The US poultry production, an industry that generates over \$469.6 billion in annual economic impact and about \$32.9 billion in taxes,[1] is organized as a private information contest. Growers – that is, independent farmers – receive chicks and feed from chicken companies, grow the chickens, and are eventually paid a fee per pound of chicken produced. These fees are based on the grower's relative performance compared to the other growers. However, growers are told neither their relative performance nor who the competing growers' are. Is such concealment good for stimulating overall poultry production? Analogous information disclosure problems arise in other settings. For instance, scholars compete for grants by submitting research projects, and presumably they ignore who they are competing against unless the grant-awarding entity publicly discloses on its webpage the list of scholars taking part in the grant competition.[2] Similarly, the organizer of a sport tournament might disclose player's information which would otherwise be too costly or impossible to retrieve by the players themselves.[3] The awareness of rival's type affects contestants' efforts: an unexceptional researcher might give up hope when told that she is competing for the grant against a leading scholar. Given this setup, I ask: "What is the optimal disclosure policy of the principal who wants to stimulate efforts?".

Running relative performance contests is a popular way to stimulate efforts that are unobservable or uncontractible. Economic theory suggests that heterogeneity of the underlying abilities of contestants is of the essence for effort exertion: unequal abilities discourage the underdog and decrease the *need* for the favorite to exert effort.[4] Empirical evidence confirms this finding.[5] Thus, a contestant's effort is

[1]Source: https://www.uspoultry.org/economic_data

[2]For such a disclosure to be effective in affecting efforts, it has to occur before the submission of the projects. In fact, project-submission contests are often structured in two stages: first, contestants submit their résumés, proofs of their eligibility or preliminary drafts of their projects, so that the contest organizer learns of and possibly discloses information about them; and second the competition takes place and contestants submit their full projects. The relevant and greatest efforts are presumably exerted in the second stage. For examples, see the German "Exzellenzinitiative", the Axa Post-Doctoral Fellowship, or the Guidelines for Architectural Design Competitions of the Royal Australian Institute of Architects. It seems customary in grant contests to not inform contestants; see the practices employed by the American NSF, by the European ERC, the British Research Councils, and the French ANR.

[3]Players in chess tournaments often ignore their rival's skills unless the tournament organizer decided to publicly disclose information about players (e.g., past performance, personal statistics, or position in overall ranking). Additionally, in recent years the NBA has disclosed on its webpage fine-tuned data collected by advanced tracking cameras. One effect of such a disclosure is that teams become better acquainted with their rival. These fine-tuned data go far beyond the commonly-known standard metrics, such as points scored by each player. For example, on stats.nba.com one can easily see that the percentage of field goals made by Danilo Gallinari when he was assisted by a teammate and his shooting position was between 25 and 29 feet from the basket in the 2014/15 season until the April 10, 2015, is 88.5%. The access to these data by teams is interpreted as a refinement on the knowledge of rival's ability.

[4]The seminal paper on promotion contests by Lazear and Rosen (1981) argues that tournament contracts with two heterogeneous workers are inefficient; the theoretical analysis by Baik (1994) discusses the adverse incentive effect of heterogeneity in a static two-player contest.

[5]Levitt (1994) shows that campaign expenditures in US House elections is highest in close races; Brown (2010) shows that the presence of a superstar in professional golf tournaments significantly lowers overall performance; Calsamiglia et al (2013) find that affirmative actions meant to equalize

affected by the information she possesses on her rival's ability. This information is often *a priori* limited, but is also refinable by an informed principal whose goal is to stimulate efforts. These features match the poultry production example and several others discussed in Section 5.

Given this setup, I ask: "What is the optimal disclosure policy of the principal?". I consider a contest between two contestants, each of whom informed of her own type and not informed of her rival's type. The principal might verifiably and costlessly inform them of their rival's type. In particular, the principal publicly announces and commits to a disclosure policy *ex-ante* – i.e., she is not yet aware of type realizations when committing to the disclosure policy.[6] Contestants' type are independent draws from a possibly *skewed* commonly-known population, or prior. The prior's skewness drives the optimal disclosure policy of the principal. For a loose intuition on how and why skewness is the driver, consider the following cases.[7]

Case I. The population consists mostly of high-types. Then a high-type strongly believes of being against another high-type even if the principal chooses to conceal types, and thus the effort of a high-type is not much affected by the principal's disclosure choice. The same is not true for a low-type: under concealment, not only is a low-type discouraged because she strongly believes her rival is a high-type, but she is *further discouraged* by knowing that her high-type rival believes she is facing another high-type and thus will exert great effort. I name this the negative Skewness Effect, or -SE, that occurs under concealment. Thus, a policy of disclosure would have the effect not only of informing a prospective low-type of (the expected news of) being up against a high, but also of informing her high rival of (the unexpected news of) being up against a low, thus disclosure would make the principal avoid the above (detrimental) further discouragement.

Case II. The population consists mostly of low-types. Then a low-type strongly believes she is up against another low-type even if the principal chooses to conceal types, and thus the effort of a low-type is not much affected by principal's disclosure choice. The same is not true for a high-type: under concealment, a high-type does not exert much effort because she strongly believes her rival is a low-type, but this negative effect on her effort is *mitigated* by knowing that her low-type rival is not discouraged, since she believes she is facing another low-type and thus will exert great effort. I name this the positive Skewness Effect, or +SE, that occurs under concealment. Thus, a policy of disclosure would have the effect not only of informing a prospective high-type of (the expected news of) being up against a low, but also of informing her low rival of (the unexpected news of) being up against a high, thus disclosure would make the principal avoid the above (beneficial) mitigation.

In Section 3, I show that under standard assumptions in contests the intuition delivered in Case I and Case II holds: aggregate effort is maximized by disclosure

contestants' opportunities are performance enhancing.

[6]This way I capture that the principal chooses the *rules* of the contest. In fact, it is seldom (if at all) observed in real-life contests that principals choose the disclosure policy at the frequency of type changes — i.e., as a function of the identity of the particular contestants. It is rather observed that principals decide and commit to an *overall* disclosure policy which is implemented every time the contest is run, like in the poultry market. Though I consider *ex-ante* disclosure, I show (in Appendix C) that if the principal could choose between an *ex-ante* and *ex-post* disclosure policy, the former is the dominant strategy.

[7]The complete formal intuition is provided in Section 3.

(concealment) when the distribution of types is skewed toward high-types (low-types). In other words, the disclosure/concealment of unexpected news is what drives the optimal disclosure policy since unexpected news is what affects efforts the most. Appendix C contains extensions of this result.

What if the principal has access to more sophisticated disclosure policies, besides those to fully disclose or fully conceal types? In Section 4, I consider a principal who has access to a technology of *partial* information disclosure, that is, who can commit to disclose or conceal *conditional* on type realizations. I show that a specific form of such partial information disclosure yields greater aggregate effort than full disclosure, full concealment, and any other partial information disclosure. In particular, the optimal partial information disclosure is to commit to disclose only the contingency which is the most favorable to the principal (i.e., contestants are all high-types), regardless of the skewness of the distribution of types. The intuition behind this result clings again to +SE and -SE, and it will be given in Section 4.

Although I make the case for the skewness effect as the driver of the optimal disclosure policy *in contests*, by no means do I attempt to make a general statement that optimal disclosure policies depend exclusively on the skewness effect. Rather, contests happen to be a suitable environment for singling out the relevance of the skewness effect, which to the best of my knowledge has received no attention so far.

Beyond what has been explained so far on the optimal disclosure policy, my paper also makes two technical contributions. First, I show a property of private information contests (Proposition 4) which generalizes a well-known wisdom of public information contests – i.e., the ratio of types equals the ratio of efforts – to private information. This property plays a crucial role in finding the optimal disclosure policy, in that it allows us to sidestep the lack of a closed-form solution of equilibrium efforts. Second, I provide a simple graphical way to single out the +SE and -SE (Appendix A). These effects, although not convoluted, has, to my knowledge, not been obtained in earlier papers. Whether the skewness effect is positive or negative relies only on the skewness of the prior and the strategic complementarity or substitutability of players' actions, not on any contest-specific feature. The characterization of SEs is thus possible in a wider class of games, although the environment we study here is restricted to contests.[8]

Several research strands are related to the present work.

1. **One-sided private information.** The prevailing fashion in the literature is to model incomplete information contests as one-sided private information on types.[9] One reason for such an approach could be that it admits a closed-

[8]For instance, Bertrand and Cournot duopoly are respectively games where strategies are strategic complements and substitutes. Thus, an elementary graphical argument analogous to the one in Appendix A.4 would show that the skewness effect is negative in the quantity or price of the high-type firm and positive in that of the low-type firm.

[9]In a one-sided private information contest the type of one contestant is common knowledge (the incumbent), and the type of the other is private information (the newcomer). The diffusion of this approach dates back to Hurley and Shogren (1998a and 1998b), and it is analyzed for example in Denter, Morgan and Sisak (2011) and Warneryd (2003). Zhang and Zhou (2015) find that in a one-sided private information setting, focusing only on full disclosure or full concealment is without loss of generality. My Theorem 2 shows that this does not carry over to two-sided private information.

form solution for equilibrium efforts, contrary to the fully private information contest.[10] In my setting it is natural to assume that a concealment policy is capable of concealing the types of each contestant, thus yielding a fully private information contest. As mentioned, Proposition 4 allows me to sidestep the lack of a closed-form solution, thus re-enabling tractability. Furthermore, the assumption of one-sided private information (as that of symmetric prior) shuts down +SE and -SE which I prove they crucially affect the optimal disclosure policy.

2. **All-pay auctions.** A parallel strand of the literature analyzes all-pay auctions rather than imperfectly discriminating contests, like (1). The all-pay auction result that is closest to my Section 3 are in Fu, Jiao and Lu (2014) and Kovenock et al. (2015). Their finding could be explained by +SE and -SE.[11]

3. **Disclosure of information not on types, and information acquisition.** Strategic information disclosure in contests other than the one analyzed here have been recently analyzed. For information disclosure of contestants' performance in a dynamic setting, see Aoyagi (2010), and of the number of contestants, see Fu et al. (2011), Lim and Matros (2009), and Myerson and Warneryd (2006). Information acquisition by contestants is analyzed by Denter, Morgan, and Sisak (2011) and Yildirim (2005) in imperfectly discriminating contests like mine, and by Morath and Münster (2013) and Szech (2011) in all-pay auctions.

4. **Bayesian persuasion.** The optimal (partial) information disclosure policy shares common features with Kamenica and Gentzkow (2015), but I will explain this relation after presenting and explaining the intuition behind my result in Section 4.

Section 2 spells out the model of contest and the disclosure game. Section 3 and Section 4 investigate the optimal disclosure policy for a principal who maximizes aggregate effort; in particular, in Section 3 the principal can only commit to either

[10] Even if the impact function $f()$ in (1) is the identity function, a contest does not admit a closed-form solution for equilibrium efforts in our fully-private information setting. Besides assuming one-sided private information, the literature has bypassed this problem of the lack of a closed-form solution by either running numerical simulations (see Hurley and Shogren, 1998a), or by modifying the standard setting to achieve a contest model which does admit a closed-form solution, such as the case of binary effort space (see Dubey, 2013), of a symmetric prior (see Malueg and Yates, 2004), or of a modified contest success function (see Wasser, 2013). My model has continuous effort space, a possibly skewed prior of types, and a standard Tullock contest success function.

[11] The -SE is due to the strategic substitutability of the effort of the low (Case I). The +SE is due to the strategic complementarity of the effort of the high (Case II). In imperfectly discriminatory contests, these two are "equally" present; that is, high- (low-) type's effort and that of her rival are strategic complements (substitutes). In all-pay auctions the contestants' efforts are strategic complements regardless of the types; that is, if a bidder bids x, her rival's best response is to bid x plus a negligible positive amount for all individually rational x's. The strategic complementarity effect is the one motivating concealment; see +SE in Case II. In fact, Fu, Jiao and Lu (2014) and Kovenock et al. (2015) find that full concealment dominates full disclosure in all-pay auctions.

fully disclose or fully conceal contestants' types, and in Section 4 the principal can also commit to partially disclose. Section 5 discusses applications and weaknesses. Appendix A provides the analysis of how beliefs about a rival's type affect effort, which is the building block of the results. Appendix B contains the proofs. Appendix C contains extentions of Theorem 1 and Theorem 2.

2 A Model of Contest

The Contest Technology. Two contestants, indexed by $i = 1, 2$, compete for a prize by exerting effort $e_i \geq 0$.[12] Each contestant has a probability of winning a prize equal to

$$p_i(e_i, e_j) = \frac{f(e_i)}{f(e_i) + f(e_j)} \tag{1}$$

with $i, j = 1, 2$, $j \neq i$. Skaperdas (1996) shows that the form (1) is the only one which satisfies a set of appealing axioms. Consider the following condition on f,

$$(\textit{Logit-CSF}) \; f(x) = x^r \text{ with } 0 < r \leq 1$$

If $e_1 = e_2 = 0$ (1) is not well-defined, but this situation is never reached in equilibrium, hence I can assume (1) takes any fixed value.

The two main results of the paper (Theorem 1 and Theorem 2) are derived under (*Logit-CSF*). Aside from being a reduced-form which has been broadly used by the literature, such an assumption is convenient for carrying out a neat analysis because of a purely technical reason which will be explained in detail at the end of the proof of Theorem 1 in Appendix B. However, the intermediate results on the role of beliefs on efforts are derived under a milder assumption (see Appendix A).

The Payoffs. Contestants are risk-neutral and compete for a prize of value normalized to 1. The cost of effort is linear, and contestant i is of type θ_i, which determines her marginal cost of effort. In particular, the payoff of a contestant of type θ_i when she exerts effort e_i and her rival exerts effort e_j is

$$p_i(e_i, e_j) - \frac{e_i}{\theta_i} \tag{2}$$

Contestant's type θ_i is an independent draw from the commonly-known prior,

$$\theta_i = \begin{cases} h & \text{with probability } p \\ l & \text{with probability } 1 - p \end{cases} \tag{3}$$

with $p \in [0, 1]$ and $h > l > 0$. Thus, being a high-type rather than a low-type brings about a lower marginal cost of effort.

[12] Restricting the number of contestants to two limits the technical difficulties due to endogenous participation, and allows me to maintain the focus of the paper on beliefs and disclosure. Moreover, it is often perceived that two is the optimal number of contestants — see, for example, Fullerton and McAfee (1999) and references therein.

Equilibrium. Since contestants are ex-ante symmetric, it is natural to focus on type-symmetric equilibria; that is, contestants of the same type follow the same equilibrium strategy, regardless of their identity.[13]

While the existence and uniqueness of equilibrium have already been proved in this setting (see Einy et al., 2016, and Ewerhart and Quartieri, 2013), interiority is not guaranteed. As shown in Appendix B, if $r \in (0,1)$, the equilibrium is interior, whereas if $r = 1$, assumption $h \leq 4l$ guarantees interiority. That is, the low-type could be sufficiently discouraged to exert 0 effort if the contest is sufficiently sensitive to efforts ($r = 1$) and high-types are more than four times stronger than low-types.

Type-symmetric equilibria and the fact that if $r = 1$ then $h \leq 4l$ are assumed throughout the paper, so as to keep the focus on disclosure and information.

The Timing of the Game. First, before types are realized, the principal commits to a disclosure policy $\mathcal{P}$, which is observed by the contestants. Then, types are realized, each contestant i learns θ_i and may or may not be informed by the principal of θ_j, according to $\mathcal{P}$. Disclosure is verifiable and costless for the principal. Finally, contestants simultaneously choose efforts. Everything – except for possibly the type realizations – is common-knowledge.

The Principal and the Disclosure Policy $\mathcal{P}$. The principal chooses $\mathcal{P}$ and maximize the expected aggregate effort. In Section 3, the principal chooses between two extreme disclosure policies on types: disclosure $\mathcal{P} = \mathcal{D}$ (that is, contestants are informed of their rival's type), and concealment $\mathcal{P} = \mathcal{C}$ (that is, contestants are not informed). In the latter case, contestants' posterior beliefs about their rival's types equal the prior (3): no information is conveyed to contestants.

In Section 4 I enlarge the space of $\mathcal{P}$'s: the principal commits to disclose or conceal *contingently* on type realizations (partial information disclosure). Thus, information might now be conveyed to contestants not only through disclosure (if any) of θ's as in Section 3, but also in case of concealment of θ's: for instance, if the principal commits to disclose only upon observing that both contestants are high-types, a high-type who has not been informed of her rival's type infers she is up against a low-type.

Each $\mathcal{P}$ induces a system of beliefs for the contestants, and each system of beliefs induces an expected level of efforts. The crucial part of the analysis is to understand how beliefs affect efforts, since then the principal simply ranks $\mathcal{P}$'s by expected efforts and commits to the best one. The analysis of how beliefs affect efforts is carried out under general contest technology – in fact, I relax (*Logit-CSF*) — and more general beliefs space than the one induced by our $\mathcal{P}$'s. This generality comes at the cost of space and it would break the flow of the paper here, hence, I store such analysis in Appendix A.

[13] More details are provided in Appendix A.

3 Optimal Disclosure Policy: Disclose or Conceal?

If the principal chooses between committing to full disclosure ($\mathcal{P} = \mathcal{D}$) or full concealment ($\mathcal{P} = \mathcal{C}$) only, which one would yield the greatest (expected) aggregate effort? When committing to a $\mathcal{P}$, the principal does not know whether the contest will be between two high-types, two low-types, or a high-type and a low-type. Table 1 provides a summary of the possible realizations of types, the corresponding probabilities, and the aggregate equilibrium effort (i.e., the principal's payoff). The notation for equilibrium efforts is as follows: the first subindex in e is the type of the contestant exerting effort, and the second subindex is the type of her rival, in case she is told (that is, under $\mathcal{D}$). Thus, for instance, e_{hl} is the equilibrium effort of a high competing against a low (under $\mathcal{D}$), and e_h is the equilibrium effort of a high who is not aware of her rival (under $\mathcal{C}$).

Realizations of $\{\theta_i, \theta_j\}$	Prob.	Aggregate effort if $\mathcal{D}$	Aggregate effort if $\mathcal{C}$
$\{h,h\}$	p^2	$2e_{hh}$	$2e_h$
$\{l,l\}$	$(1-p)^2$	$2e_{ll}$	$2e_l$
$\{h,l\}$ or $\{l,h\}$	$2p(1-p)$	$e_{hl} + e_{lh}$	$e_h + e_l$

Table 1: Possible realizations of contestants' types, with corresponding probability and aggregate equilibrium efforts under $\mathcal{P} = \mathcal{D}$ and $\mathcal{P} = \mathcal{C}$.

The more skewed is the prior (p far from 0.5), the more likely are contingencies of evenness of types ($\{h,h\}$ and $\{l,l\}$) types to be symmetric, and in line with the conventional wisdom that an unevenness of types reduces efforts. and following the contest literature wisdom. Thus, one might conjecture that

Throughout this section I gradually build the intuition that will eventually lead to Theorem 1, starting with how beliefs affect efforts – in this section, beliefs under $\mathcal{C}$ simply coincide with the prior (3).

Figure 1 plots equilibrium efforts under $\mathcal{D}$ (the four thick solid lines) and under $\mathcal{C}$ (the two thin solid lines) as functions of p.[14] These functions allow a clear visualization of +SE and -SE — which are intuitively explained in Case I and Case II of the Introduction — and thus help understanding the main result of this section.

Efforts under $\mathcal{D}$.

1. $e_{hh} > e_{hl}$ and $e_{ll} > e_{lh}$. A contestant exerts her greatest effort when competing against a contestant of the same type (even contest), because competition is maximum. When she is instead competing against a stronger contestant, she is discouraged, and when competing against a weaker contestant, she does not need much effort to be sufficiently confident of winning (uneven contest). This is the conventional wisdom that an unevenness of types reduces efforts.

[14]Plots are created using the polynomial in e_h-only (or e_l-only) obtained from the simplification of (4) and (5) with $f(x) = x$, $l = 1$ and $h = 2$. The qualitative features of the figures do not depend on these latter assumptions.

2. e_{hh}, e_{hl}, e_{lh}, and e_{ll} do not depend on p. Type disclosure is verifiable, thus under disclosure the prior distribution of types does not affect contestants' efforts.[15]

Efforts under $\mathcal{C}$.

1. Focus first on features of e_h.

   (a) e_h is increasing in p, because p increases the probability of an even contest.

   (b) $\lim_{p\to 1} e_h = e_{hh}$. When $p \to 1$, a high is basically aware of being against another high if $p \to 1$.

   (c) $\lim_{p\to 0} e_h > e_{hl}$. One might think that, symmetrical to the previous point, $\lim_{p\to 0} e_h = e_{hl}$, however, this does not happen; in fact, the reasoning of a high-type under $p \to 0$ is, *"I essentially know that I am against a low, but my low rival thinks she is competing against another low as $p \to 0$, and therefore she will not give up hope and will exert great effort."* That is, the high-type best replies to e_{ll} rather than to e_{lh}, and $e_{ll} > e_{lh}$. Hence, since the effort of the high-type and that of her rival are strategic complements, the best reply to e_{ll} is greater than the one to e_{lh}.[16] This causes the upward jump called +SE in Figure 1 (see Case II in the Introduction). The +SE fades away as p moves away from 0 because it becomes less likely that high and low have such unequal beliefs of being in an even contest like the one leading to +SE (when $p = \frac{1}{2}$ high and low have equal beliefs of being in an even contest).

2. Features of e_l are analogous to those of e_h, and in particular $\lim_{p\to 1} e_l < e_{lh}$. In fact, under $\mathcal{C}$ and $p \to 1$, a low exerts *less* effort than if she was up against a high under $\mathcal{D}$ because her high rival thinks that she is in an even contest – therefore exerting e_{hh} rather than e_{hl}. Hence, since the effort of the low-type and that of her rival are strategic substitutes, the best reply to e_{hh} is lower than the best reply to e_{hl}. This causes the downward jump called –SE in Figure 1 (see Case I in the Introduction), and it fades away as p moves away from 1.

With the above understanding in mind, the intuition behind the optimal choice of $\mathcal{P}$ is easily delivered. If two highs or two lows will realize (first and second line of Table 1), $\mathcal{D}$ maximizes aggregate effort because it prevents contestants from thinking that they are competing in an uneven contest (i.e., $2e_{hh} \geq 2e_h$ and $2e_{ll} \geq 2e_l$). If instead a high and a low will realize (third line of Table 1), $\mathcal{C}$ maximizes aggregate

[15] The fact that $e_{hl} > e_{lh}$ is not needed, although it is a feature of contests under (1). In fact, the reader can consider Figure 1 separately for efforts of the high-type (e_{hh}, e_{hl}, e_h) and those of the low-type (e_{ll}, e_{lh}, e_l). In my model the principal has no control over type-realizations.

[16] There is no jump in e_h as $p \to 1$ because a high believes she is up against a high, who believes she is up against another high, and so on.

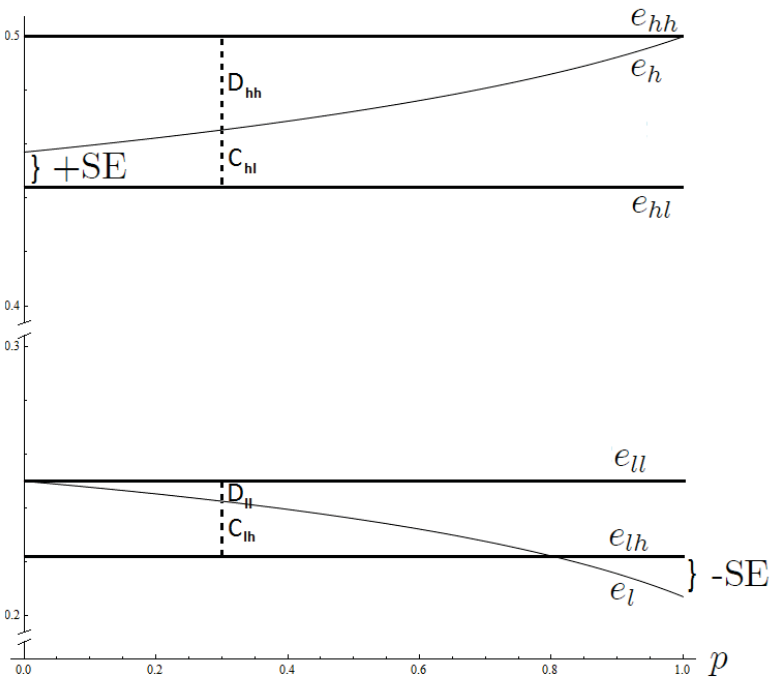

Figure 1: Equilibrium efforts as functions of p assuming $r = 1$, $l = 1$ and $h = 2$. The thick lines are the efforts under $\mathcal{D}$ (e_{hh}, e_{hl}, e_{lh}, and e_{ll}). The thin lines are the efforts under $\mathcal{C}$ (e_h and e_l). The two segments named +SE and -SE are respectively the positive and negative skewness effects. The dotted lines are the incentive to commit to $\mathcal{D}$ or $\mathcal{C}$ for $p = 0.3$.

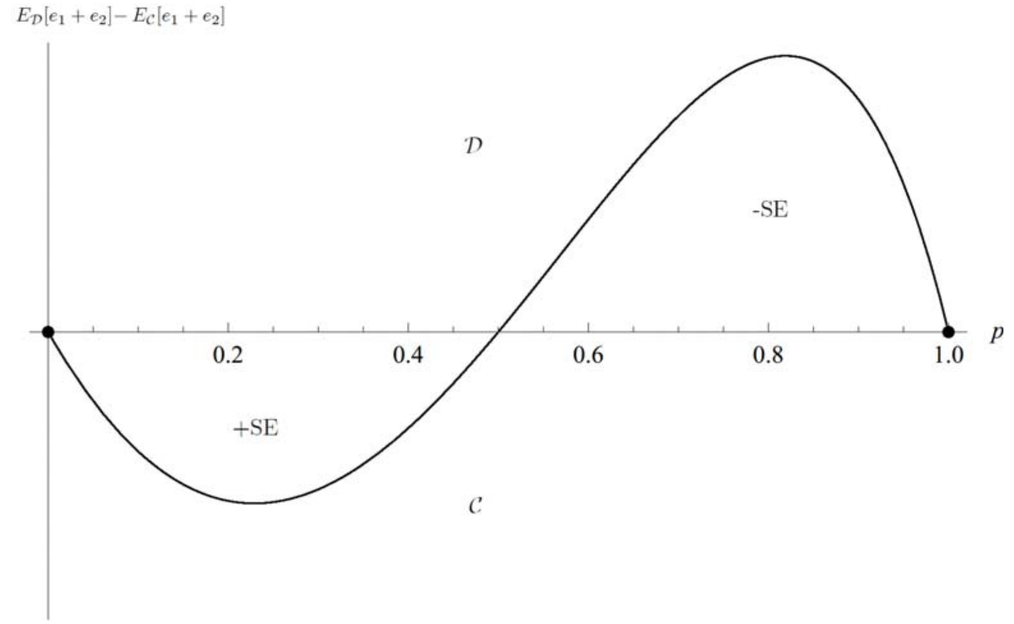

Figure 2: $E_{\mathcal{D}}[e_1+e_2]-E_{\mathcal{C}}[e_1+e_2]$: Expected aggregate effort under $\mathcal{D}$ minus expected aggregate effort under $\mathcal{C}$ as a function of p, assuming $r=1$, $l=1$ and $h=2$.

effort because it hides the unevenness of types. Therefore, finding the optimal $\mathcal{P}$ boils down to the *ex-ante* trade-off between the benefits of $\mathcal{P}=\mathcal{D}$ if the contest will be even (realizations $\{h,h\}$ and $\{l,l\}$) and the benefits of $\mathcal{P}=\mathcal{C}$ if the contest will be uneven (realizations $\{h,l\}$ and $\{l,h\}$). Let me visualize these two opposite benefits in Figure 1 as vertical dotted lines (depicted for $p=0.3$). The benefits of $\mathcal{P}=\mathcal{D}$ if $\{h,h\}$ (resp., $\{l,l\}$) is that each of them exerts e_{hh} (resp., e_{ll}) rather than e_h (resp., e_l); this is D_{hh} (resp., D_{ll}) in Figure 1. The benefits of $\mathcal{P}=\mathcal{C}$ if $\{h,l\}$ or $\{l,h\}$ is that the high exerts e_h rather than e_{hl} and the low exerts e_l rather than e_{lh}; this is C_{hl} and C_{lh} in Figure 1. Here is where +SE and -SE enter the game. The upward jump in e_h due to +SE for low p's shrinks D_{hh} and inflates C_{hl}. The downward jump in e_l due to -SE for high p's inflates D_{ll} and shrinks C_{lh}. Therefore, +SE and -SE motivate $\mathcal{D}$ for sufficiently low p's and $\mathcal{C}$ for sufficiently high p's. These results lead to the optimal $\mathcal{P}$ of Theorem 1, which is depicted in Figure 2.[17] When $p=\frac{1}{2}$, the +SE and the -SE balance out (see Proposition 9 in Appendix A), because uncertainty on the rival's type is both maximum and symmetric across types. This is a case of what I call a SE-unbiased contest in Appendix A.

Theorem 1 *Under (Logit-CSF) and the possibility of full disclosure or full concealment only, the optimal disclosure policy for the principal is to commit to full concealment (full disclosure) if the distribution of types is skewed toward low- (high-) types, that is, if* $0\leq p\leq\frac{1}{2}$ ($\frac{1}{2}\leq p\leq 1$). *The principal is indifferent between full*

[17] Although parameters h and l do not affect the principal's optimal disclosure policy, the absolute value of the difference between aggregate effort under $\mathcal{D}$ and under $\mathcal{C}$ increases in the ratio h/l. In other words, the vertical scale of Figure 2 shrinks for a reduction of h/l, all the way down to the case of $h=l$, that is, a flat horizontal line depicting indifference between any $\mathcal{P}$ for all values of p.

concealment and full disclosure if and only if the prior is degenerate, i.e., $p \in \{0, 1\}$, or symmetric, i.e., $p = \frac{1}{2}$.

The only trivial part of Theorem 1 is that if $p \in \{0, 1\}$ information disclosure plays no role and the principal is indifferent between disclosure and concealment – that is, $\mathcal{P}$ will not affect efforts. Appendix C contains extensions of Theorem 1 under continuum type space (*C.1*) and under correlated types (*C.2*). Also, it is shown that an ex-ante choice of $\mathcal{P}$ dominates an ex-post choice of $\mathcal{P}$ – that is, the principal observes types and then chooses $\mathcal{C}/\mathcal{D}$ (*C.3*). Finally, alternative objective functions for the principal are discussed (*C.4*).

4 Optimal Disclosure Policy: Partial Information Disclosure

In Section 3, the principal could only commit to fully disclose or fully conceal types. Can the principal do any better using a partial disclosure policy? This section shows that the answer is positive when allowing for *type-contingent* disclosure: in particular, a specific form of type-contingent disclosure dominates full disclosure, full concealment, and any other type-contingent disclosure, regardless of p.

The principal can now commit to disclose or conceal *contingently* on the realization of types. That is, $\mathcal{P}$ is now a vector of three binary variables, each taking the value $\mathcal{D}$ or $\mathcal{C}$, where the first (respectively, second and third) element corresponds to the disclosure choice in case of contingency $\{h, h\}$ (respectively, $\{h, l\}$ and $\{l, l\}$).[18] For instance, under disclosure policy $\mathcal{P} = \{\mathcal{D}, \mathcal{C}, \mathcal{C}\}$, the only contingency in which contestants are told their rival's type is when both contestants are high-types. As in Section 3, $\mathcal{P}$ is publicly announced ex-ante by the principal who commits to it.[19] The result of this section is that, in fact, $\mathcal{P} = \{\mathcal{D}, \mathcal{C}, \mathcal{C}\}$ turns out to be optimal.

Theorem 2 *Under (Logit-CSF) and the possibility of partial information disclosure, the optimal disclosure policy for the principal is to commit to only disclose realization $\{h, h\}$. That is, $\mathcal{P} = \{\mathcal{D}, \mathcal{C}, \mathcal{C}\}$.*

[18] Note that I abuse notation by keeping the same $\mathcal{P}$ in Section 3 and Section 4, although they are defined over different spaces. Also, contingency $\{h, l\}$ and contingency $\{l, h\}$ are equivalent for the principal, thus in any optimal $\mathcal{P}$ there is no incentive to assign different disclosure choices to $\{h, l\}$ and $\{l, h\}$.

[19] Although the disclosure policy analyzed here formally nests the two extreme cases analyzed in Section 3 - $\mathcal{P} = \{\mathcal{C}, \mathcal{C}, \mathcal{C}\}$ corresponds to $\mathcal{P} = \{\mathcal{C}\}$ in Section 3, and $\mathcal{P} = \{\mathcal{D}, \mathcal{D}, \mathcal{D}\}$ corresponds to $\mathcal{P} = \{\mathcal{D}\}$ in Section 3 -, I keep the two sections separate because: i) the comparison of public and private information contests is of self-interest, and ii) the more sophisticated disclosure policy in this section might not be implementable in some real-life situations since it requires a strong commitment power by the principal.

Moreover, under partial information disclosure, the comparison of expected aggregate effort with and without commitment is trivial; commitment dominates. The non-trivial comparison I relegate to *C.3* in Appendix C is between: i) commitment to full disclosure or to full concealment, and ii) lack of commitment, or equivalently ex-post disclosure.

The intuition behind Theorem 2 clings to +SE and -SE, as the one behind Theorem 1. First consider how $\mathcal{P} = \{\mathcal{D}, \mathcal{C}, \mathcal{C}\}$ affects contestants' beliefs. A high-type who is not told her rival's type infers she must be up against a low-type. Thus, (A) a high always knows her rival's type, and (B) a low never knows her rival's type. Even more importantly, (A) and (B) are common-knowledge; low-types know (A) and high-types know (B). Therefore, such a $\mathcal{P}$ allows the principal to both:

- avoid the losses due to -SE. -SE is an effort loss, occurring when a low faces a high not who is not aware she is facing a low (see Case I). This situation *does not* occur under such a $\mathcal{P}$; even if a low faces a high, the low knows that the high is aware that she is facing a low. That is, low-types know (A).

- benefit from +SE. +SE is an effort gain, occurring when a high faces a low who is unaware she is facing a high (see Case II). This situation *does* occur under such a $\mathcal{P}$; a high-type knows that, when facing a low, this low will not be aware that she is facing a high. That is, high-types know (B) above.

Though +SE and -SE are crucial for understanding the above intuition, Theorem 2 holds regardless of the skewness itself – that is, $\forall p$. This suggests that when skewness is toward low-types, the optimal policy relies on making the most out of the +SE, and when skewness is toward high-types, the optimal policy relies on avoiding the -SE.

The practice suggested by Theorem 2 provides a rationale for some existing practice in contests. Say a firm opens a top-position to both current employees (internal contestants) and to the public (external contestants). Arguably, the internal contestants have a *competitive advantage* over the external contestants; they have skills and experiences more in line with the firm's requirements, they are already experienced in the firm's practices, or they perhaps know some members of the selection committee personally. Aside from the competitive advantage, the internal contestants also have an *informational advantage* over the external contestants; they are more likely to have access to the list of contestants or simply to know whether any other internal is taking part. Without much of a stretch, this example rationalizes informed-high-types (internal candidates) and uninformed-low-types (external candidates), as in Theorem 2. The same logic applies to a contest for a grant where some scholars might have both a competitive and informative advantage by having a friend in the grant committee; a competitive advantage because of having a better grasp of what the committee is looking for, besides perhaps having the playing-field levelled toward them by their friend, and an informational advantage because of information leakages by their friend on the list of competing scholars. However, the information structure of these two applications is exogenous, rather than being the outcome of the principal's choice. Yet, in several contests it is arguably easy to implement the policy of informing only the top-players, which Theorem 2 suggests. For instance, in the poultry industry example, the optimal partial information disclosure might take the form of disclosing the set of contestants only to "top-growers"; that is, for instance, growers whose historical productivity is above average.

The optimal partial information disclosure policy of Theorem 2 shares common features with the bayesian persuasion literature started with Kamenica and Gentzkow (2015), henceforth KG. Consider the introductory example of KG. A defendant is guilty or innocent with some commonly-known prior probabilities. A prosecutor (sender) tries to convince a judge (receiver) that the defendant is guilty.

The judge wants to choose the just action (convict if guilty, acquit if innocent). The prosecutor commits to reporting to the judge a signal on the defendant's type, which he will observe by conducting an investigation. KG show that it is optimal for the prosecutor to commit to the following signal structure: if he observes the defendant's type that is favorable to him (guilty), he discloses it, and if he observes that the defendant is innocent he strategically randomizes between reporting innocent and reporting guilty. That way, the judge who receives the report guilty does not know whether it comes from the true state of the world being guilty or from the randomization in case of innocence. In other words, the sender's optimal report is to commit to disclose the best signal for him (that is, guilty) and pool all other signals together in a strategic way. Likewise, in my setting, the optimal (partial information) disclosure for the sender (principal) is to commit to disclose the best signal for him (that is, contingency $\{h, h\}$) and pool all other signals together (that is, conceal them). The main difference between KG and this paper lies in the way the pooling of the "non-best" signals is implemented: in KG, it is assumed that the sender cannot conceal but can commit to a stochastic disclosure policy; in the present paper, it is assumed that the sender cannot stochastically disclose but can commit to conceal.

5 Discussion

The linkage principle proposed by Milgrom and Weber (1982) suggests that the revenue-maximizing auctioneer should commit *ex-ante* to fully disclose to the bidders all available information. The present paper can be seen as a step ahead in addressing the same question in contests. Theorem 1 and Theorem 2 entail that the linkage principle does not carry over to contests. In particular, Theorem 1 states that full disclosure extracts more (less) expected aggregate effort than full concealment if the distribution of types is skewed toward high-types (low-types). Theorem 2 states that if partial information disclosure is possible, the expected aggregate effort is maximum in case of committing to disclose only the signal which is best for the principal – that is, contestants are all high-types.

These two theorems provide neat testable implications on levels of efforts under different information regimes. Admittedly, the intuition behind these results requires advanced reasoning. Thus, evidence in favor of these results might suggest a high level of sophistication of the players, whereas evidence in contrast with these results might suggest that players fail to achieve the second-level reasoning behind +SE and -SE, which drives the results.[20]

A number of shortcomings of my analysis pinpoints room for future research. First, in many of the discussed applications it is sensible to say that the goal of the principal is to stimulate aggregate effort. For this reason, I assumed the principal maximizes aggregate effort, as is commonly assumed in the literature on contests. However, one could imagine a principal with a different goal, such as the effort exerted by the winner which I analyze in *C.4* in the Appendix, or others which I leave

[20]For instance, if $E_{\mathcal{D}}[e_1 + e_2] - E_{\mathcal{C}}[e_1 + e_2]$ in Figure 2 is found to be not statistically different from the horizontal axis, it suggests that +SE and -SE are not taken into account by players.

for future research. Second, I ignored contestants' selection and participation issues. For the former, it may be effort-improving to endow the model with an entry fee capable of sifting out weak applicants. For the latter, contestants' participation in the contest may depend on the declared disclosure policy itself, and I conjecture that concealment deters the participation of the low-types, and that this might be beneficial in terms of expected aggregate effort. Third, the disclosure of information about other contestants that would otherwise be ignored by contestants has several other consequences aside from the one considered in the present paper – that is, of directly affecting contestants' efforts. In fact, disclosure may also trigger communication among contestants,[21] affect the external visibility of contestants,[22] and create animosity among contestants.[23] I abstracted away from these issues and focused here on the direct effects on efforts of making contestants' types publicly available to other contestants. Fourth, I assumed contestants can be of one of two types only, high or low. This greatly simplified the analysis; in equilibrium, the strategy of the high-type is a strategic complement, as is the strategy of the low substitute, *regardless* of their beliefs about their rival's type. Strategic complementarity/substitutability is a key driver of the results as it allows us to tell apart +SE and -SE. If the type-space had, say, a medium-type, her effort would be a strategic complement or substitute for her rival's effort depending on her beliefs about the rival's type; strategic substitutability if she sufficiently believes she is up against a high rather than a low, and complementarity otherwise. This conditional strategic substitutability/complementarity for the medium type would add a substantial layer of technical difficulty to the analysis and I leave it to future research. For a special tractable case of the continuum of types, see *C.1* in the Appendix.

[21] Communication among contestants might have both negative and positive effects on competition: negative as it facilitates collusion and positive as it facilitates a better understanding of the problem at hand (for public procurement, "The purpose of exchanging information is to improve the understanding of Government requirements and industry capabilities, thereby allowing potential offerors to judge whether or how they can satisfy the Government's requirements, and enhancing the Government's ability to obtain quality supplies and services, including construction, at reasonable prices, and increase efficiency in proposal preparation, proposal evaluation, negotiation, and contract award", see FAR 15.201 (b))

[22] The external visibility could be positive (fostering contestants' visibility in the market, or preventing illegal lobbying activities) or negative (jeopardizing contestants' anonymity or exposing them to media pressure).

[23] Disclosing the list of competing employees and their relative ranking might put a strain on them and might hurt overall productivity, morale, and teamwork — see Milkovich and Newman (1996).

A Appendix A. Intermediate Results: The Role of Beliefs on Efforts

In this appendix I derive results which are instrumental in achieving Theorem 1 and Theorem 2. In particular, the proof of Theorem 1 relies on Proposition 4 (discussed in *A.2*, and proved in *B.2*), and the proof of Theorem 2 further relies on Proposition 5 (discussed in *A.3*, and proved in *B.3*). Moreover, in *A.4* I focus on +SE and -SE; in particular, first I visualize them in the graph of best replies, second I analytically show their existence (Corollary 6), and third I discuss the conditions for a contest to be what I call "SE-unbiased" (Corollary 8 and Proposition 9).

A.1 Preliminaries

The results of this appendix are derived under greater *generality* than stated in the main text. Unfortunately, this generality cannot be kept until the final results (Theorem 1 and Theorem 2); the reason is technical and I explain it in Appendix B.8. The generality of this appendix is twofold; on the contest technology (1), and on the $\mathcal{P}$-induced beliefs.

Regarding the *contest technology*, throughout this appendix and unless otherwise stated, the following is assumed

$$(Reg)\ f(0) = 0,\ f'(\cdot) > 0 \text{ and } f''(\cdot) \leq 0$$

(Reg) guarantees the existence and uniqueness of equilibrium in pure strategy.[24] Also, under (Reg) FOCs are necessary and sufficient to characterize the best reply, which is continuously differentiable and bounded.[25]

Regarding the *$\mathcal{P}$-induced beliefs*, the generality of this appendix aims to include possibly asymmetric information released to contestants through a feasible $\mathcal{P}$, so as to embrace the feasible beliefs of both Section 3 and Section 4. Thus, beliefs are a probability distribution over types: p_h (p_l) denotes the belief of a high (low) of being in an even contest, that is, of being against another high (low). More formally,

$$\begin{aligned} p_h &= E_{\mathcal{P}}[\theta_j = h | \theta_i = h] \\ p_l &= E_{\mathcal{P}}[\theta_j = l | \theta_i = l] \end{aligned}$$

with $i, j = 1, 2$, $j \neq i$. For equilibrium consistency, beliefs are common-knowledge. Note that under $\mathcal{P} = \mathcal{C}$, $p_h = p$ and $p_l = 1 - p$.

As said, I focus on type-symmetric strategies. That is, all contestants of type θ_i follow the same equilibrium strategy e_{θ_i}, regardless of their identity. Therefore, the system of FOCs has only two unknowns (e_h and e_l) and two beliefs parameters (p_h

[24]See Szidarovszky and Okuguchi (1997) for the existence and uniqueness under complete information. Under incomplete information, Einy et al. (2013) and Ewerhart and Quartieri (2013) prove existence and uniqueness in a setting that nests my model.

[25]See Yildirim (2005) and Morgan and Várdy (2007).

and p_l).[26]

$$(\text{FOC of } h\text{-type}): \qquad p_h \frac{f'(e_h)}{4f(e_h)} + (1-p_h)\frac{f'(e_h)f(e_l)}{[f(e_h)+f(e_l)]^2} = \frac{1}{h} \tag{4}$$

$$(\text{FOC of } l\text{-type}): \qquad p_l \frac{f'(e_l)}{4f(e_l)} + (1-p_l)\frac{f'(e_l)f(e_h)}{[f(e_h)+f(e_l)]^2} = \frac{1}{l} \tag{5}$$

where f' denotes the derivative of f. SOCs hold by (*Reg*). An equilibrium is a pair (e_h, e_l) satisfying (4) and (5).

A.2 Proposition 4

If the principal ends up *disclosing* then $(p_h, p_l) \in \{0,1\}^2$ because contestants are certain of their rival's type, regardless of whether the disclosure comes from a full or a partial disclosure policy. Then, the equilibrium efforts under (Logit-CSF) are a well-known result (see for example Nti, 1999) easily retrievable from the system of (4) and (5):

Lemma 3 *Under $\mathcal{D}$ and (Logit-CSF), equilibrium efforts are given by*

$$e_{hh} = \frac{rh}{4}, \quad e_{hl} = \frac{rh^{r+1}l^r}{(h^r+l^r)^2}, \quad e_{lh} = \frac{rl^{r+1}h^r}{(h^r+l^r)^2}, \quad e_{ll} = \frac{rl}{4} \tag{6}$$

If the principal ends up *concealing* then (p_h, p_l) can be interior and depend on whether the concealment comes from a full or partial concealing policy. For example, if concealment is observed under $\mathcal{P} = \{\mathcal{D}, \mathcal{C}, \mathcal{C}\}$ then $p_h = 0$ and $p_l = 1-p$. In the special case of observing concealment as a result of a full concealment policy, then contestants' prior and posterior coincide since $\mathcal{P}$ does not convey any information; that is, $p_h = p$ and $p_l = 1-p$. Nevertheless, the system (4) and (5) generally lacks a closed-form solution for (e_h, e_l).[27] In fact, the very existence of a (pure-strategies) equilibrium was not established until recently (see Einy et al., 2016, and Ewerhart and Quartieri, 2013). However, simple manipulations of the system of (4) and (5) lead to Proposition 4, which allows us to sidestep the need for a closed-form solution, by providing a property which is sufficient to characterize the optimal disclosure policy.

Proposition 4 *The following holds:*

$$\frac{f'(e_{ll})}{f(e_{ll})}\frac{f(e_l)}{f'(e_l)}(1-p_h) = \frac{f'(e_{hh})}{f(e_{hh})}\frac{f(e_h)}{f'(e_h)}(1-p_l) + (p_l - p_h) \tag{7}$$

In particular, under (Logit-CSF),

$$\frac{e_l}{e_{ll}}(1-p_h) = \frac{e_h}{e_{hh}}(1-p_l) + (p_l - p_h) \tag{8}$$

[26] To lighten the notation I omit throughout the paper the dependencies of e_h and e_l on p_h, p_l, h, l.

[27] Even if $p_h = p$ and $p_l = 1-p$ and f being the identity function, there is no closed-form solution for (e_h, e_l) unless $p \in \{0, \frac{1}{2}, 1\}$. See Malueg and Yates (2008) for equilibrium effort in case $p = 1/2$.

Proposition 4 provides a neat relation between the equilibrium efforts under $\mathcal{C}$ (i.e., e_h and e_l) and the equilibrium efforts under $\mathcal{D}$ and symmetric types (i.e., e_{hh} and e_{ll}), where the notation for the subindexes of equilibrium efforts is the one introduced at the beginning of Section 3.

More importantly, Proposition 4 helps find the optimal disclosure policy. Consider Figure 2. Proposition 4 will allow me to prove that there are exactly three crossings with the horizontal axis (in 0, $\frac{1}{2}$, and 1) and that the derivative in $p = \frac{1}{2}$ is strictly negative. By continuity, the shape depicted in Figure 2 will follow, characterizing the optimal disclosure policy without need for closed-form equilibrium efforts. This is the structure of the proof of Theorem 1 in *B.7*.

A.3 Proposition 5

To see how beliefs affect effort, let us first see how beliefs affect best replies.[28] I start with extreme values of p_h and p_l. If $p_h = 1$ then (4) alone identifies e_h – which equals e_{hh} –, and the best reply is inelastic to e_l (see Figure 3). Symmetrically, if $p_l = 1$ then (5) alone identifies e_l – which equals e_{ll} –, and the best reply is inelastic to e_h (see Figure 4). If $p_h = p_l = 0$ then (4) and (5) interact, the best replies have maximum elasticity (see Figure 3 and 4 with $p_h = 0$ and $p_l = 0$), and they cross in $e_h = e_{hl}$ and $e_l = e_{lh}$.[29] Thus, increasing p_h and p_l flattens out the best replies of h and l respectively, and a contestant's effort is maximum when she (thinks she) is competing with another contestant of the same type (the inelastic best replies in Figure 3 and Figure 4). This is the conventional wisdom that the unevenness of types reduces individual efforts, re-interpreted from the belief viewpoint. I store these facts in i) and ii) of the following proposition.

Proposition 5 *Efforts increase in the beliefs of being in an even contest, regardless of types: that is, i)* $\frac{\partial e_h}{\partial p_h} > 0$*, and ii)* $\frac{\partial e_l}{\partial p_l} > 0$*.*

Effort of the high (low) increase (decrease) in the low's (high's) belief of being in an even contest: that is, iii) $\frac{\partial e_h}{\partial p_l} > 0$*, and iv)* $\frac{\partial e_l}{\partial p_h} < 0$*.*

However, the effort of a high is affected not only by her beliefs, but also by the beliefs of the low, and viceversa: results iii) and iv) are the by-products of such reasoning and, respectively, the strategic complementarity of the effort of the high and the strategic substitutability of the effort of the low. That is, what I rename SEs. An increase in p_l (p_h), flattens out the best reply of the low (high) who exerts more effort because of her increased belief of being in an even contest, and this translates into an upright-shift (downright-shift) of the crossing of the best replies because the high (low) responds by increasing (decreasing) her effort.

[28] With a lexical abuse, I name best reply, respectively of the high and of the low, the set of points in the (e_h, e_l)-space satisfying respectively (4) and (5).

[29] Graphical understanding of the role of strategic complementarity/substitutability in contests through the use of best reply functions has had a long-standing tradition since the seminal contribution of Dixit (1987). His graphical analysis corresponds in my case to $p_h = p_l = 0$.

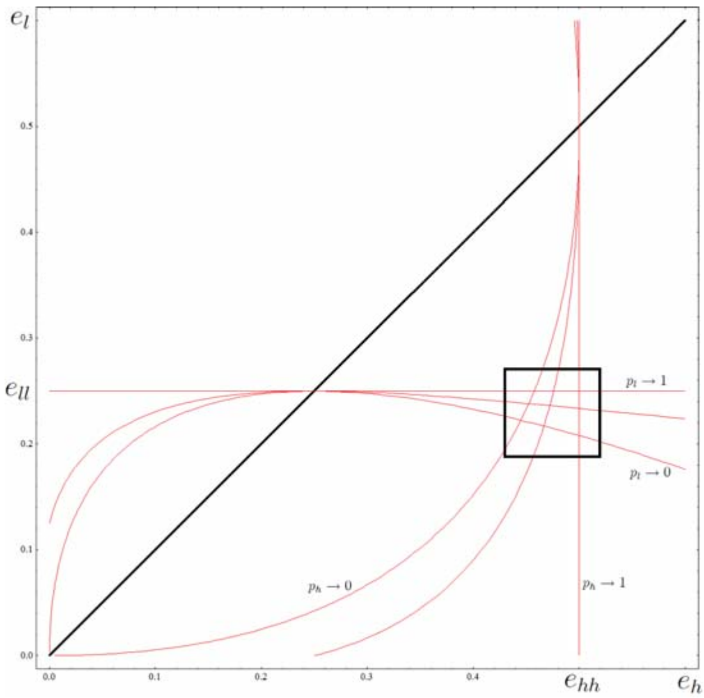

Figure 3: Plot of Equation (4) and Equation (5): best reply of a high and best reply of a low as their beliefs vary, with $f(x) = x$, $h = 2$, and $l = 1$. The region in the rectangle will be zoomed in Figure 4.

A.4 Skewness Effects: +SE and -SE

+SE and -SE have geometrical counterparts in the graph of best replies of Figure 3. To see this, Figure 4 zooms into the region in the rectangle of Figure 3 and shows how to visualize +SE and -SE as geometrical distances.

Analytical existence and signs of the skewness effects can be easily proved in the limits as a corollary of Proposition 5.

Corollary 6 *i)* $\lim_{p\to 0} e_h > e_{hl}$, *ii)* $\lim_{p\to 0} e_l = e_{ll}$, *iii)* $\lim_{p\to 1} e_h = e_{hh}$, *and iv)* $\lim_{p\to 1} e_l < e_{lh}$. *In particular, +SE can be seen in i), and -SE can be seen in iv).*

In the remainder of this appendix I analyze the conditions on beliefs for a contest to be *not* affected by skewness effects. It is clear at this point that the skewness effects exist only when beliefs of being in an even contest are different for high-types and low-types. Thus, to rule out skewness effects I need that high-types believe they are up against another high-type with a certain probability and low-types believe they are up against another low-type with the *same* probability. I define such a situation an SE-unbiased contest.

Definition 7 *An **SE-unbiased contest** is a contest where* $p_h = p_l$.

For instance, $p = 1/2$ and full concealment yields an SE-unbiased contest.[30] Two properties of SE-unbiased contests are worth highlighting.

[30]Correlated types also implies SE-unbiased contest, as I discuss in Appendix C.

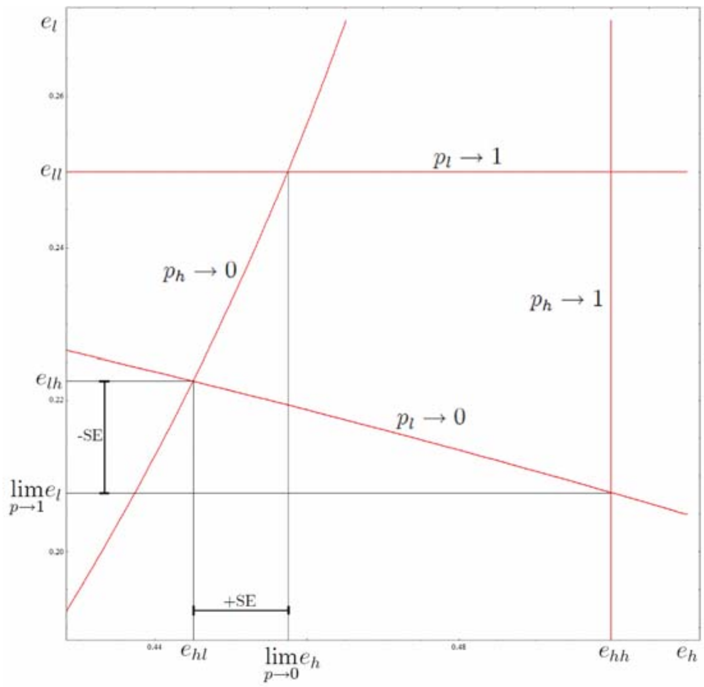

Figure 4: Region of Figure 3 in the rectangle, and visualization of +SE and -SE. Full concealment implies $p_h = p$ and $p_l = 1-p$ (see the prior (3)). If $p_h \to 1$, a high believes she is in an even contest, and hence $e_h \to e_{hh}$ regardless of p_l. If $p_l \to 1$, a low believes she is in an even contest, and hence $e_l \to e_{ll}$ regardless of p_l. If $p_h \to 0$ and $p_l \to 0$, high and low believe they are in an uneven contest under complete information, and hence play e_{hl} and e_{lh} respectively (this case resembles Dixit, 1987). If $p_h \to 0$ and $p_l \to 1$, then $e_h \to e_{hl} + (+SE)$ since a high believes she is up against a low who believes she is facing another low. These latter beliefs are generated by full concealment and $p \to 1$. If $p_h \to 1$ and $p_l \to 0$, then $e_l \to e_{lh} + (-SE)$ since a low believes she is up against a high who believes she is facing another high. These latter beliefs are generated by full concealment and $p \to 0$.

First, a straightforward consequence of Proposition 4 is that,

Corollary 8 *The following holds:*

$$p_h \lesseqqgtr p_l \iff \frac{f(e_h)}{hf'(e_h)} \lesseqqgtr \frac{f(e_l)}{lf'(e_l)} \tag{9}$$

In particular, under (Logit-CSF),

$$p_h \lesseqqgtr p_l \iff \frac{e_h}{h} \lesseqqgtr \frac{e_l}{l} \tag{10}$$

Under (*Logit-CSF*) and full information (that is, $\mathcal{D}$) it is well-known that the ratio of efforts equals the ratio of contestants' types. This can be seen in (10): a high and a low competing in a full information contest have beliefs such that $p_h = p_l = 0$, thus $e_h/h = e_l/l$.[31] Hence, not only does Corollary 8 extend to a general impact function f the fact that the ratio of efforts equals the ratio of types (9), but it also proves that under (*Logit-CSF*) this fact carries over to imperfect information contests if and only if the contest is SE-unbiased. Equality of ratios of types and ratio of efforts is useful because it allows us to write the system of the two FOCs as a single equation in one unknown only, which is then solvable either explicitly or implicitly.

Second,

Proposition 9 *In an SE-unbiased contest with (Logit-CSF),*

i) $e_h = pe_{hh} + (1-p)e_{hl}$

ii) $e_l = pe_{lh} + (1-p)e_{ll}$

iii) expected aggregate effort is equal in $\mathcal{D}$ *and* $\mathcal{C}$.

Proposition 9 highlights the convenience and drawback of assuming a symmetric prior (which implies SE-unbiased contest). The convenience is to have a closed-form solution for $\mathcal{C}$-equilibrium efforts, which I lack in my model with a possibly skewed prior. The drawback is that a symmetric prior shuts down the skewness effects – which are not present in Malueg and Yates' work – which I showed to drive differences in the expected aggregate effort under different disclosure policies.

Malueg and Yates (2004) propose a model which is *de facto* an SE-unbiased contest: in their prior, h and l are equally and symmetrically correlated.[32] In fact, they also find the results in Proposition 9.[33]

[31] Proposition 2 in Corchón (2000) proves that this result holds under homogeneity of degree zero of the contest success function (together with mild regularity assumptions).

[32] The probability that if a contestant is h also the other contestant is h equals the probability that if a contestant is l also the other contestant is l. Thus, in my notation, $p_l = p_h$.

[33] The first two results are (26) in their work, whereas the third result is their Proposition 5.

B Appendix B. Proofs

The existence of a (unique) equilibrium is a special case of Theorem 1 in Einy et al. (2016), and of Theorem 1 in Ewerhart and Quartieri (2013). In this Appendix, first, I provide preliminary results which simplify the subsequent proofs (Lemma 10-14), then I discuss the condition for interiority, and finally I prove Proposition 4 and Proposition 5.

Lemma 10 $\frac{f'(x)}{f(x)}$ *is strictly decreasing in* x

Proof. $\frac{\partial}{\partial x}\left[\frac{f'(x)}{f(x)}\right] = \frac{f''(x)f(x)-[f'(x)]^2}{[f(x)]^2} < 0$ follows from (*Reg*). ■

Lemma 11 $e_h > e_l$.

Proof. Rewrite the system of (4) and (5) as

$$\begin{cases} p_h A + (1-p_h)B = \frac{1}{h} \\ p_l C + (1-p_l)D = \frac{1}{l} \end{cases} \tag{11}$$

where I define

$$A \equiv \frac{f'(e_h)}{4f(e_h)},\ B \equiv \frac{f'(e_h)f(e_l)}{[f(e_h)+f(e_l)]^2},\ C \equiv \frac{f'(e_l)}{4f(e_l)},\ D \equiv \frac{f'(e_l)f(e_h)}{[f(e_h)+f(e_l)]^2}$$

Assume by contradiction that $e_h \leq e_l$. Since $\frac{1}{h} < \frac{1}{l}$, (11) implies

$$p_h A + (1-p_h)B < p_l C + (1-p_l)D \tag{12}$$

Also, it is routine to show that $A \geq B$, $A \geq C$ (by Lemma 10) and $C \geq D$ and $B \geq D$ (by Lemma 10). Thus, $A \geq \max\{B, C\} \geq \min\{B, C\} \geq D$. If $B \geq C$, then a contradiction is immediately reached by $A \geq B \geq C \geq D$ and (12). If instead $C > B$, a contradiction will be reached in the remainder of this proof.

$$\begin{aligned} C &> B \\ \frac{f'(e_l)}{4f(e_l)} &> \frac{f'(e_h)f(e_l)}{[f(e_h)+f(e_l)]^2} \\ [f(e_h)+f(e_l)]^2 &> 4[f(e_l)]^2 \frac{f'(e_h)}{f'(e_l)} \end{aligned} \tag{13}$$

By $f''(\cdot) \leq 0$ and $e_h \leq e_l$, it follows that $f'(e_l) \leq f'(e_h)$, and hence it is necessary for (13) that

$$\begin{aligned} [f(e_h)+f(e_l)]^2 &> 4[f(e_l)]^2 \\ [f(e_h)]^2 + 2f(e_h)f(e_l) - 3[f(e_l)]^2 &> 0 \\ [f(e_h)+3f(e_l)][f(e_h)-f(e_l)] &> 0 \end{aligned}$$

which contradicts $e_h \leq e_l$.Hence, $e_h > e_l$. ■

Lemma 12 *The left-hand side of (4) decreases in* e_h *and increases in* e_l.

Proof. Using the notation of Lemma 11, A decreases in e_h by Lemma 10. B decreases in e_h because its numerator decreases in e_h and its denominator increases in e_h. A is constant in e_l. B increases in e_l because by Lemma 11

$$\begin{aligned}\frac{\partial B}{\partial e_l} &= f'(e_h)\frac{f'(e_l)\left[f(e_h)+f(e_l)\right]-2f(e_l)f'(e_l)}{\left[f(e_h)+f(e_l)\right]^3} \\ &= \frac{f'(e_h)f'(e_l)}{\left[f(e_h)+f(e_l)\right]^3}\left[f(e_h)-f(e_l)\right]>0\end{aligned}$$

■

Lemma 13 *The left-hand side of (5) decreases in both e_h and e_l*

Proof. The proof is analogous to that of Lemma 12 and it is omitted here. ■

Lemma 14 $\frac{f(e_h)}{f'(e_h)} \leq \frac{h}{4}$ *and* $\frac{f(e_l)}{f'(e_l)} \leq \frac{l}{4}$.

Proof. Consider (4). If $p_h = 1$, $\frac{f(e_h)}{f'(e_h)} = \frac{h}{4}$. Consider now the effect of lowering $p_h(<1)$. Then, by $A \geq B$ (using the notation of Lemma 11), the convex combination on the left-hand side of (4) decreases in p_h. To keep it equal to the constant $\frac{1}{h}$, then e_h must decrease because both A and B are decreasing functions of e_h (see the proof of Lemma 12). Hence, $\frac{f(e_h)}{f'(e_h)} \leq \frac{h}{4}$. The proof of $\frac{f(e_l)}{f'(e_l)} \leq \frac{l}{4}$ is symmetric. ■

Interiority of Equilibrium

If any corner, it must be e_l, and cannot be $e_h = 0$, since it would contradict Lemma (11). When $e_l = 0$? By Proposition 5, e_l is the lowest when $p_h = 1$ and $p_l = 0$, which into (4) and (5) yields the following system of FOCs,

$$\frac{f'(e_h)}{4f(e_h)} = \frac{1}{h} \tag{14}$$

$$\frac{f'(e_l)f(e_h)}{\left[f(e_h)+f(e_l)\right]^2} = \frac{1}{l} \tag{15}$$

(14) has e_{hh} as the unique solution:

$$f(e_{hh}) = \frac{h}{4}f'(e_{hh}) \tag{16}$$

in particular, under (*Logit-CSF*), $e_{hh} = \frac{rh}{4}$.

The left-hand side of (15) is decreasing in e_l by Lemma 13, hence $e_l > 0$ if

$$\begin{aligned}\frac{f'(0)f(e_{hh})}{\left[f(e_{hh})+f(0)\right]^2} &> \frac{1}{l} \\ \frac{f'(0)}{f(e_{hh})} &> \frac{1}{l} \\ \frac{f'(0)}{f'(e_{hh})} &> \frac{h}{4l} \end{aligned} \tag{17}$$

where the first step is due to $f(0) = 0$ and the second step is due to (16). Condition (17) guarantees interiority under (*Reg*). In the special case of (*Logit-CSF*), as discussed in Section 2, condition (17) holds:

- if $r \in (0, 1)$, because its left-hand side tends to infinity,
- if $r = 1$ and I further assume that $h \leq 4l$.

Proof of Proposition 4

In (4) and (5) isolate the second addends of the left-hand sides and divide them to obtain

$$\frac{(1-p_h)f'(e_h)f(e_l)}{(1-p_l)f'(e_l)f(e_h)} = \frac{\frac{1}{h} - p_h\frac{f'(e_h)}{4f(e_h)}}{\frac{1}{l} - p_l\frac{f'(e_l)}{4f(e_l)}}$$

Cross-multiply to obtain

$$\begin{aligned}
\left(\frac{1}{l} - p_l\frac{f'(e_l)}{4f(e_l)}\right)(1-p_h)f'(e_h)f(e_l) &= \left(\frac{1}{h} - p_h\frac{f'(e_h)}{4f(e_h)}\right)(1-p_l)f'(e_l)f(e_h) \\
\frac{f'(e_l)}{4f(e_l)}\left(\frac{4}{l}\frac{f(e_l)}{f'(e_l)} - p_l\right)(1-p_h)f'(e_h)f(e_l) &= \frac{f'(e_h)}{4f(e_h)}\left(\frac{4}{h}\frac{f(e_h)}{f'(e_h)} - p_h\right)(1-p_l)f'(e_l)f(e_h) \\
\left(\frac{4}{l}\frac{f(e_l)}{f'(e_l)} - p_l\right)(1-p_h) &= \left(\frac{4}{h}\frac{f(e_h)}{f'(e_h)} - p_h\right)(1-p_l) \\
\frac{4}{l}\frac{f(e_l)}{f'(e_l)}(1-p_h) &= \frac{4}{h}\frac{f(e_h)}{f'(e_h)}(1-p_l) + (p_l - p_h) \qquad (18)
\end{aligned}$$

Also, e_{hh} and e_{ll} are the solution of (4) and (5) with $p_h = p_l = 1$. That is,

$$\frac{f'(e_{ll})}{f(e_{ll})} = \frac{4}{l} \qquad (19)$$

$$\frac{f'(e_{hh})}{f(e_{hh})} = \frac{4}{h} \qquad (20)$$

(19) and (20) into (18) yields (7). Finally, under (*Logit-CSF*), $\frac{f'(x)}{f(x)} = \frac{r}{x}$ and (8) follows.

Proof of Proposition 5

(7) is equivalent to

$$\frac{\frac{4}{l}\frac{f(e_l)}{f'(e_l)} - p_l}{1-p_l} = \frac{\frac{4}{h}\frac{f(e_h)}{f'(e_h)} - p_h}{1-p_h} \qquad (21)$$

and remember that both e_l and e_h depend on (p_h, p_l).

[Proof of Part iv) $\partial e_l/\partial p_h < 0$] Assume by contradiction that $\partial e_l/\partial p_h \geq 0$. Consider the left-hand side of (21); it increases in e_l by Lemma 10, and thus by $\partial e_l/\partial p_h \geq 0$ it also increases in p_h. Since the left-hand side of (21) increases in p_h, also the right-hand side of (21) has to increase in p_h. However, $\frac{x-p_h}{1-p_h}$ decreases in p_h whenever $x \leq 1$ (which holds by Lemma 10), hence the only way to have the right-hand side of (21) increasing in p_h is that $\partial e_h/\partial p_h \geq 0$. By Lemma 13, $\partial e_l/\partial p_h \geq 0$ and $\partial e_h/\partial p_h \geq 0$ lead to a contradiction. Therefore, $\partial e_l/\partial p_h < 0$.

[Proof of Part i) $\partial e_h/\partial p_h > 0$] An increase in p_h decreases e_l (as just proved), and hence by Lemma 13 to keep the left-hand side of (5) constantly equal to $1/l$, e_h must increase in p_h.

[Proof of Part ii) $\partial e_l/\partial p_l > 0$] e'_h denotes $\partial e_h/\partial p_l$, and e'_l denotes $\partial e_l/\partial p_l$. By Lemma 12, from the left-hand side of (4) I can retrieve that $e'_l > 0$ iff $e'_h > 0$. I will now prove that $e'_l \leq 0$ and $e'_h \leq 0$ lead to a contradiction. I do so by differentiating (5) with respect to p_l (for the differential of $(1-p_h)B$ I apply the formula in (36))

$$\frac{\overbrace{[f'(e_l)+p_l f''(e_l)e'_l]\,f(e_l)}^{E}\overbrace{-p_l\,[f'(e_l)]^2\,e'_l}^{F}}{4\,[f(e_l)]^2}+$$
$$+\frac{\overbrace{-f'(e_l)f(e_h)}^{G}\overbrace{+(1-p_l)\,[f''(e_l)f(e_h)e'_l+f'(e_l)f'(e_h)e'_h]}^{H}}{[f(e_h)+f(e_l)]^2}+$$
$$\overbrace{-\frac{2(1-p_l)f'(e_l)f(e_h)\,[f'(e_h)e'_h+f'(e_l)e'_l]}{[f(e_h)+f(e_l)]^3}}^{I}=0$$

I prove that $E+F+G+H+I>0$ to achieve a contradiction and thus end the proof.

Term F is trivially positive.[34]

$$\begin{aligned}
E+G &= \frac{f'(e_l)f(e_l)\,[f(e_h)+f(e_l)]^2-4f'(e_l)f(e_h)\,[f(e_l)]^2+p_l f''(e_l)f(e_l)e'_l\,[f(e_h)+f(e_l)]^2}{4\,[f(e_l)]^2\,[f(e_h)+f(e_l)]^2}\\
&= \frac{f'(e_l)f(e_l)\,[f(e_h)-f(e_l)]^2+p_l f''(e_l)f(e_l)e'_l\,[f(e_h)+f(e_l)]^2}{4\,[f(e_l)]^2\,[f(e_h)+f(e_l)]^2}>0
\end{aligned}$$

$$\begin{aligned}
H+I &= (1-p_l)\frac{[f''(e_l)f(e_h)e'_l+f'(e_l)f'(e_h)e'_h]\,[f(e_h)+f(e_l)]-2f'(e_l)f(e_h)\,[f'(e_h)e'_h+f'(e_l)e'_l]}{[f(e_h)+f(e_l)]^3}\\
&\geq (1-p_l)\frac{f'(e_l)f'(e_h)e'_h\,[f(e_h)+f(e_l)]-2f'(e_l)f(e_h)f'(e_h)e'_h}{[f(e_h)+f(e_l)]^3}\\
&= (1-p_l)f'(e_l)f'(e_h)e'_h\frac{f(e_l)-f(e_h)}{[f(e_h)+f(e_l)]^3}>0
\end{aligned}$$

where the last inequality holds true by Lemma 11. Therefore, $e'_l \leq 0$ and $e'_h \leq 0$ lead to a contradiction and the result follows.

[Proof of Part iii) $\partial e_h/\partial p_l > 0$] It directly follows from $\partial e_l/\partial p_l > 0$ and the fact that $\partial e_l/\partial p_l > 0$ iff $\partial e_l/\partial p_l > 0$ (see the proof of Part ii) above).

Proof of Corollary 6

Consider the system of (4) and (5).

i) $\lim_{p\to 0} e_h$ is the solution of the system under $p_h \to 0$ and $p_l \to 1$.

e_{hl} is the solution of the system under $p_h \to 0$ and $p_l \to 0$.

Thus, by iii) of Proposition 5, $\lim_{p\to 0} e_h > e_{hl}$.

ii) $\lim_{p\to 0} e_l$ is the solution of the system under $p_h \to 0$ and $p_l \to 1$.

[34] Remember that we are under assumption $e'_l \leq 0$ and $e'_h \leq 0$.

e_{ll} is the solution of the system under $p_l \to 1$ and $\forall p_h$ (including $p_h \to 0$). The result follows.

iii) $\lim_{p\to 1} e_h$ is the solution of the system under $p_h \to 1$ and $p_l \to 0$.

e_{hh} is the solution of the system under $p_h \to 1$ and $\forall p_l$ (including $p_l \to 0$). The result follows.

iv) $\lim_{p\to 1} e_l$ is the solution of the system under $p_h \to 1$ and $p_l \to 0$.

e_{lh} is the solution of the system under $p_h \to 0$ and $p_l \to 0$.

Thus, by iv) of Proposition 5, $\lim_{p\to 1} e_l < e_{lh}$.

Proof of Proposition 9

The proof mirrors the steps of Malueg and Yates (2004) to get to their expression (25) and to their Proposition 6, and hence it is omitted.

Proof of Theorem 1

I adopt the compact notation $\pi^{\mathcal{D}-\mathcal{C}}$ for $E_{\mathcal{D}}[e_1 + e_2] - E_{\mathcal{C}}[e_1 + e_2]$; that is, the difference in the expected sum of efforts under $\mathcal{D}$ and under $\mathcal{C}$, which is depicted in Figure 2 as a function of p. I prove the claim in three steps. **Step 1:** $\pi^{\mathcal{D}-\mathcal{C}} = 0$ iff $p \in \{0, \frac{1}{2}, 1\}$. **Step 2:** The derivative of $\pi^{\mathcal{D}-\mathcal{C}}$ with respect to p in $p = \frac{1}{2}$ is strictly positive. **Step 3:** $\pi^{\mathcal{D}-\mathcal{C}}$ is continuous in p. These three results together lead to the sign of $\pi^{\mathcal{D}-\mathcal{C}}$ as in Figure 2, and hence Theorem 1 follows.

Step 1. I show that $\pi^{\mathcal{D}-\mathcal{C}} = 0$ iff $p \in \{0, \frac{1}{2}, 1\}$. First, I analyze when $\pi^{\mathcal{D}-\mathcal{C}}$ takes the value 0.

$$\pi^{\mathcal{D}-\mathcal{C}} = p^2[2e_{hh} - 2e_h] + 2p(1-p)[e_{hl} + e_{lh} - e_h - e_l] + (1-p)^2[2e_{ll} - 2e_l] = 0$$

i.e.,

$$p^2 e_{hh} + (1-p)^2 e_{ll} + p(1-p)[e_{hl} + e_{lh}] - pe_h - (1-p)e_l = 0 \tag{22}$$

If $p = 0$ or $p = 1$, by $e_l = e_{ll}$ and $e_h = e_{hh}$ respectively, (22) holds and $\pi^{\mathcal{D}-\mathcal{C}} = 0$. Thus, from now on I can focus on $p \in (0, 1)$.

which upon substitution of (6) and (8) yields

$$\begin{aligned}
p^2\frac{rh}{4} + (1-p)^2\frac{rl}{4} + p(1-p)\frac{r(h+l)h^r l^r}{(h^r + l^r)^2} - p\frac{h+l}{h}e_h - \frac{1-2p}{4}rl &= 0 \\
p^2\frac{rh}{4} + p^2\frac{rl}{4} + p(1-p)\frac{r(h+l)h^r l^r}{(h^r + l^r)^2} - p\frac{h+l}{h}e_h &= 0 \\
pr\frac{h+l}{4} + (1-p)\frac{r(h+l)h^r l^r}{(h^r + l^r)^2} &= \frac{h+l}{h}e_h \\
pe_{hh} + (1-p)e_{hl} &= e_h \qquad (23)
\end{aligned}$$

Hence, (23) is a condition for the indifference between $\mathcal{D}$ and $\mathcal{C}$ written in terms of the efforts exerted by the high-type only. With a similar procedure used to find (23) – that is, by substituting (8) into e_h rather than into e_l – I can obtain the value of e_l for which the administrator is indifferent between $\mathcal{D}$ and $\mathcal{C}$, which symmetrically to (23) is

$$(1-p)e_{ll} + pe_{lh} = e_l \tag{24}$$

Plug (23) and (24) into the top equation of (4) under (*Logit-CSF*) and see if any $p \in (0,1)$ solves the resulting equation – thus, yielding $\pi^{\mathcal{D}-\mathcal{C}} = 0$.[35]

First, use (6) to rewrite the indifference conditions (23) and (24) for e_l and e_h as

$$e_h = rh\frac{p(h^r + l^r)^2 + 4(1-p)h^r l^r}{4(h^r + l^r)^2} \tag{25}$$

$$e_l = rl\frac{(1-p)(h^r + l^r)^2 + 4ph^r l^r}{4(h^r + l^r)^2} \tag{26}$$

These efforts are those that, if exerted under full concealment, lead to indifference between $\mathcal{C}$ and $\mathcal{D}$. Now, using (4) and (5), I check whether these effort levels are reached for some parameter values. Hence, I rewrite (4) as

$$pr\frac{h}{4} + (1-p)r\frac{he_h^r e_l^r}{(e_h^r + e_l^r)^2} = e_h \tag{27}$$

Plugging (25) into the right-hand side of (27), and after simple simplifications, I obtain the following

$$\frac{e_h^r e_l^r}{(e_h^r + e_l^r)^2} = \frac{h^r l^r}{(h^r + l^r)^2} \tag{28}$$

Finally, I plug (25) and (26) where I defined $J = p(h^r + l^r)^2 + 4(1-p)h^r l^r$ and $K = (1-p)(h^r + l^r)^2 + 4ph^r l^r$ into (28), and obtain

$$\begin{aligned}
\frac{h^r l^r J^r K^r}{(h^r J^r + l^r K^r)^2} &= \frac{h^r l^r}{(h^r + l^r)^2} \\
h^{2r} J^r K^r + l^{2r} J^r K^r + 2h^r l^r J^r K^r &= h^{2r} J^{2r} + l^{2r} K^{2r} + 2h^r l^r J^r K^r \\
l^{2r} K^r (J^r - K^r) &= h^{2r} J^r (J^r - K^r) \qquad (29)
\end{aligned}$$

and the unique solution of (29) is $J = K$,[36] which is equivalent to

$$\begin{aligned}
p(h^r + l^r)^2 + 4(1-p)h^r l^r &= (1-p)(h^r + l^r)^2 + 4ph^r l^r \\
4(1-2p)h^r l^r &= (1-2p)(h^r + l^r)^2
\end{aligned}$$

whose unique solution is $p = \frac{1}{2}$. Similar algebra shows that (25), (26) and $p = \frac{1}{2}$ satisfy (5) – besides satisfying (4) as proved. Hence, I proved that there are only three values of p for which $\pi^{\mathcal{D}-\mathcal{C}} = 0$: 0, $\frac{1}{2}$, and 1.

Step 2. I write the system (4) and (5) as a unique equation in terms of e_h and parameters only, and then I make use of the implicit function theorem to evaluate the derivative of $\pi^{\mathcal{D}-\mathcal{C}}$ in $p = \frac{1}{2}$, and prove that it is strictly positive. That is,

$$\left.\frac{\partial \pi^{\mathcal{D}-\mathcal{C}}}{\partial p}\right|_{p=\frac{1}{2}} > 0$$

[35] Remark: the fact that (23) and (24) are sufficient for $\pi^{\mathcal{D}-\mathcal{C}} = 0$ could have already been noticed in (22), but I needed to use (8) to show that they are also necessary for $\pi^{\mathcal{D}-\mathcal{C}} = 0$.

[36] Note that $J = K \iff \frac{e_h}{e_l} = \frac{h}{l}$. See Proposition 4 and Corollary 8.

Remember that efforts under $\mathcal{D}$ are not functions of p, unlike the efforts under $\mathcal{C}$. I omitted this detail so far in the notation, and I now write it when it would otherwise yield confusion as I need to differentiate with respect to p.

To simplify $\pi^{\mathcal{D}-\mathcal{C}}$ I use the same steps used to move from (22) to (23), where I simplified a p, and get

$$\begin{aligned}
\frac{\partial}{\partial p}\left[p^2 e_{hh} + p(1-p)e_{hl} - pe_h(p)\right]\Bigg|_{p=\frac{1}{2}} &> 0 \\
\left[2pe_{hh} + e_{hl} - 2pe_{hl} - e_h(p) - p\frac{\partial e_h(p)}{\partial p}\right]\Bigg|_{p=\frac{1}{2}} &> 0 \\
e_{hh} - e_h\left(\frac{1}{2}\right) &> \frac{1}{2}\left(\frac{\partial e_h(p)}{\partial p}\bigg|_{p=\frac{1}{2}}\right) \qquad (30)
\end{aligned}$$

When $p = \frac{1}{2}$, I know from Step 1 that $\pi^{\mathcal{D}-\mathcal{C}} = 0$, and hence from (23), $e_h\left(\frac{1}{2}\right) = \frac{e_{hh}+e_{hl}}{2}$. Therefore, (30) is equivalent to

$$e_{hh} - e_{hl} > \frac{\partial e_h(p)}{\partial p}\bigg|_{p=\frac{1}{2}} \qquad (31)$$

The left-hand side of (31) is known by (6). The right-hand side is trickier. First, isolate e_l in (8):

$$\begin{aligned}
(1-p)\frac{e_l}{e_{ll}} &= p\frac{e_h}{e_{hh}} + (1-2p) \\
(1-p)e_l &= p\frac{l}{h}e_h + (1-2p)\frac{rl}{4} \\
e_l &= \frac{4ple_h + (1-2p)rlh}{4(1-p)h} \qquad (32)
\end{aligned}$$

Use (32) into (4), and obtain

$$f(e_h, p) \equiv p\frac{r}{4e_h} + 4^r h^r l^r r\frac{(1-p)^{r+1}e_h^{r-1}[h(1-2p) + 4pe_h]^r}{[4^r h^r (1-p)^r e_h^r + (4ple_h + hl(1-2p))^r]^2} - \frac{1}{h} = 0$$

The defined $f(e_h, p)$ is an equation in p and e_h only, and hence by the implicit function theorem

$$\frac{\partial e_h(p)}{\partial p}\bigg|_{p=\frac{1}{2}} = -\frac{\frac{\partial f(e_h,p)}{\partial p}\Big|_{p=\frac{1}{2}}}{\frac{\partial f(e_h,p)}{\partial e_h}\Big|_{p=\frac{1}{2}}} \qquad (33)$$

I will eventually plug (33) into (31) to conclude the proof of Step 2. I start with the denominator of (33):

$$
\begin{aligned}
\left.\frac{\partial f(e_h,p)}{\partial e_h}\right|_{p=\frac{1}{2}} &= \frac{\partial f(e_h,\frac{1}{2})}{\partial e_h} \\
&= \left.\frac{\partial}{\partial e_h}\left[\frac{r}{8e_h} + 2^{2r}h^r l^r r \frac{2^{-(r+1)}e_h^{r-1}\,(2e_h)^r}{[2^r h^r e_h^r + 2^r l^r e_h^r]^2}\right]\right|_{p=\frac{1}{2}} \\
&= \left.\frac{\partial}{\partial e_h}\left[\frac{r}{8e_h} + \frac{h^r l^r r}{2\,(h^r+l^r)^2\,e_h}\right]\right|_{p=\frac{1}{2}} \\
&= \left.\left[-\frac{r}{8e_h^2} - \frac{h^r l^r r}{2(h^r+l^r)^2 e_h^2}\right]\right|_{p=\frac{1}{2}} \\
&= -\frac{r}{e_h^2}\left.\left[\frac{h^{2r}+l^{2r}+6h^r l^r}{8(h^r+l^r)^2}\right]\right|_{p=\frac{1}{2}} \qquad (34)
\end{aligned}
$$

Note that when $p=\frac{1}{2}$ the contest is what I called SE-unbiased, and its equilibrium effort is $e_h = rh\frac{h^{2r}+l^{2r}+6h^r l^r}{8(h^r+l^r)^2}$ (see Proposition 9). I use this fact into (34) and obtain

$$
\left.\frac{\partial f(e_h,p)}{\partial e_h}\right|_{p=\frac{1}{2}} = -\left.\frac{1}{he_h}\right|_{p=\frac{1}{2}}
$$

Hence, expression (33) reads

$$
\begin{aligned}
\left.\frac{\partial e_h(p)}{\partial p}\right|_{p=\frac{1}{2}} &= -\frac{\left.\frac{\partial f(e_h,p)}{\partial p}\right|_{p=\frac{1}{2}}}{\left.\frac{\partial f(e_h,p)}{\partial e_h}\right|_{p=\frac{1}{2}}} \\
&= \left(he_h\left.\frac{\partial f(e_h,p)}{\partial p}\right|_{p=\frac{1}{2}}\right) \\
&= \frac{rh}{4} + 4^r h^{r+1} l^r r e_h^r\,\frac{\partial}{\partial p}\left.\left[\frac{(1-p)^{r+1}[h(1-2p)+4pe_h]^r}{[4^r h^r(1-p)^r e_h^r + (4ple_h + hl(1-2p))^r]^2}\right]\right|_{p=\frac{1}{2}} \\
&= \frac{rh}{4} + 4^r h^{r+1} l^r r e_h^r\,\frac{\partial}{\partial p}\left.\left[\frac{a(p)b(p)}{[c(p)]^2}\right]\right|_{p=\frac{1}{2}} \qquad (35)
\end{aligned}
$$

where I defined

$$
\begin{aligned}
a(p) &= (1-p)^{r+1} \\
b(p) &= [h(1-2p)+4pe_h]^r \\
c(p) &= 4^r h^r (1-p)^r e_h^r + (4ple_h + hl(1-2p))^r
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\frac{\partial}{\partial p}\left.\left[\frac{a(p)b(p)}{[c(p)]^2}\right]\right|_{p=\frac{1}{2}} &= \left.\frac{[a'(p)b(p)+a(p)b'(p)]\,[c(p)]^2 - 2a(p)b(p)c(p)c'(p)}{[c(p)]^4}\right|_{p=\frac{1}{2}} \\
&= \left.\frac{a'(p)b(p)+a(p)b'(p)}{[c(p)]^2}\right|_{p=\frac{1}{2}} - \left.\frac{2a(p)b(p)c'(p)}{[c(p)]^3}\right|_{p=\frac{1}{2}} \qquad (36)
\end{aligned}
$$

From the definitions of the functions a, b and c compute their values and their derivatives when $p = 1/2$.

$$\begin{array}{ll}
a\left(\frac{1}{2}\right) = \frac{1}{2^{r+1}} & a'\left(\frac{1}{2}\right) = -\frac{r+1}{2^r} \\
b\left(\frac{1}{2}\right) = 2^r e_h^r & b'\left(\frac{1}{2}\right) = 2^{r+1} r e_h^{r-1}[e_h - h/2] \\
c\left(\frac{1}{2}\right) = 2^r e_h^r (h^r + l^r) & c'\left(\frac{1}{2}\right) = -2^{r+1} h^r r e_h^r + 2^{r+1} l^r r e_h^{r-1}(e_h - h/2)
\end{array}$$

Plug these results into (36) to write (35) in the following way[37]

$$\begin{aligned}
\left.\frac{\partial e_h(p)}{\partial p}\right|_{p=\frac{1}{2}} &= \frac{rh}{4} - 4^r h^{r+1} l^r r e_h^r \left[\frac{e_h^r(r+1) - r e_h^{r-1}(e_h - h/2)}{\left[2^r e_h^r (h^r + l^r)\right]^2} + \frac{r e_h^r 2^{r+1}\left[-h^r e_h^r + l^r e_h^{r-1}(e_h - h/2)\right]}{\left[2^r e_h^r (h^r + l^r)\right]^3} \right] \\
&= \frac{rh}{4} - h^{r+1} l^r r \frac{1 + r e_h^{-1}(h/2)}{\left[(h^r + l^r)\right]^2} - 2h^{r+1} l^r r^2 \frac{-h^r + l^r - l^r e_h^{-1}(h/2)}{(h^r + l^r)^3} \\
&= \frac{rh}{4} + h^{r+2} l^r r^2 \frac{2l^r - (h^r + l^r)}{2e_h (h^r + l^r)^3} - h^{r+1} l^r r \frac{l^r(2r+1) + h^r(1-2r)}{(h^r + l^r)^3} \\
&= \frac{rh}{4} + 4h^{r+1} l^r r \frac{l^r - h^r}{(h^r + l^r)\left[(h^r + l^r)^2 + 4h^r l^r\right]} - h^{r+1} l^r r \frac{l^r(2r+1) + h^r(1-2r)}{(h^r + l^r)^3}
\end{aligned}$$

where in the last step I used the fact that $e_h = rh\frac{h^{2r} + l^{2r} + 6h^r l^r}{8(h^r + l^r)^2}$ when $p = \frac{1}{2}$. Therefore, I can finally evaluate expression (31).

$$\begin{aligned}
\frac{rh}{4} - rh\frac{h^r l^r}{(h^r + l^r)^2} &> \frac{rh}{4} + 4h^{r+1} l^r r \frac{l^r - h^r}{(h^r + l^r)\left[(h^r + l^r)^2 + 4h^r l^r\right]} + \\
&\quad - h^{r+1} l^r r \frac{l^r(2r+1) + h^r(1-2r)}{(h^r + l^r)^3} \\
\frac{l^r(2r+1) + h^r(1-2r) - h^r - l^r}{(h^r + l^r)^2} &> 4\frac{l^r - h^r}{\left[(h^r + l^r)^2 + 4h^r l^r\right]} \\
2r\frac{l^r - h^r}{(h^r + l^r)^2} &> 4\frac{l^r - h^r}{\left[(h^r + l^r)^2 + 4h^r l^r\right]} \\
2(h^r + l^r)^2 &> r\left[(h^r + l^r)^2 + 4h^r l^r\right]
\end{aligned}$$

By $r \leq 1$, it suffices to show that

$$\begin{aligned}
2(h^r + l^r)^2 &> (h^r + l^r)^2 + 4h^r l^r \\
(h^r - l^r)^2 &> 0
\end{aligned}$$

and the result follows.

Step 3. The continuity of e_h and e_l in p directly follows from the Maximum Theorem applied to contestants' payoff, which is continuous and strictly concave in own effort. The continuity of $\pi^{\mathcal{D}-\mathcal{C}}$ in p follows from the continuity of e_h and e_l.

Why is *(Logit-CSF)* convenient?

Conditions $pe_{hh} + (1-p)e_{hl} = e_h$ and $(1-p)e_{ll} + pe_{lh} = e_l$ are the conditions for indifference between $\mathcal{C}$ and $\mathcal{D}$ – respectively called (23) and (24) in the above.

[37] For the sake of brevity, I use e_h rather than $e_h|_{p=\frac{1}{2}}$.

Imagine to draw their left-hand sides in Figure 1. They are straight lines in p going from e_{hl} to e_{hh} the former, and from e_{ll} to e_{lh} the latter. Where and how many times do they cross e_h the former and e_l the latter? Thanks to assumption (*Logit-CSF*),

(A) e_h is concave in p, and thus there is a unique value of p for which $pe_{hh} + (1-p)e_{hl} = e_h$ holds.[38]

(B) e_l is convex in p, and thus there is a unique value of p for which $(1-p)e_{ll} + pe_{lh} = e_l$ holds.

(C) these two unique values of p both occur in $p = 1/2$.[39]

While (*Logit-CSF*) guarantees (A)-(B)-(C), (*Reg*) alone guarantees none of them, and thus it undermines the fact that $\pi^{\mathcal{D}-\mathcal{C}} = 0$ in a unique point (i.e., in $p = 1/2$). Can we find an assumption milder than (*Logit-CSF*) and stronger than (*Reg*) which would still make $\pi^{\mathcal{D}-\mathcal{C}} = 0$ for a unique p – perhaps different than 1/2? Guaranteeing (A)-(B) could be easily done with a mild condition on f'''.[40] Nevertheless, without (C), uniqueness of p for which $\pi^{\mathcal{D}-\mathcal{C}} = 0$ is not granted, and $\pi^{\mathcal{D}-\mathcal{C}}$ might cross the horizontal axis several times. In fact, if say $p = 0.3$ satisfies $pe_{hh} + (1-p)e_{hl} = e_h$ and $p = 0.7$ satisfies $(1-p)e_{ll} + pe_{lh} = e_l$, then there could be multiple $p \in [0.3, 0.7]$ such that $\pi^{\mathcal{D}-\mathcal{C}}$ changes sign.

B.1 Proof of Theorem 2

The notation of beliefs – p_h and p_l – carries over here, thus if for example $\mathcal{P} = \{\mathcal{D}, \mathcal{C}, \mathcal{C}\}$ and contestants observe $\mathcal{C}$, then $p_h = 0$ and $p_l = 1 - p$. In other words, it needs not to be the case anymore that $p_h = 1 - p_l$ as in Section 3. Denote by $\pi^{\mathcal{P}}$ the expected sum of efforts under disclosure policy $\mathcal{P}$. First, I prove that $\pi^{\{\mathcal{D},\mathcal{D},\mathcal{D}\}} > \pi^{\{\mathcal{C},\mathcal{C},\mathcal{D}\}}$ (**Lemma 15**), second that $\pi^{\{\mathcal{D},\mathcal{C},\mathcal{C}\}} > \pi^{\{\mathcal{D},\mathcal{D},\mathcal{D}\}}$ (**Lemma 16**), and third that $\pi^{\{\mathcal{D},\mathcal{C},\mathcal{C}\}} > \pi^{\{\mathcal{C},\mathcal{C},\mathcal{C}\}}$ (**Lemma 17**). Fourth, note that in all the remaining policies (i.e., $\{\mathcal{C},\mathcal{D},\mathcal{D}\}, \{\mathcal{D},\mathcal{C},\mathcal{D}\}, \{\mathcal{D},\mathcal{D},\mathcal{C}\}, \{\mathcal{C},\mathcal{D},\mathcal{C}\}$) contestants perfectly infer types, and thus they are outcome-equivalent to $\mathcal{P} = \{\mathcal{D},\mathcal{D},\mathcal{D}\}$. Theorem 2 follows.

The proofs of lemmas 15, 16 and 17 are alike. First, I simplify the difference in the expected sum of efforts under the two disclosure policies using (21) with the appropriate p_l and p_h. Second, I use the results in Proposition 5 to conclude the proof.

Lemma 15 $\pi^{\{\mathcal{D},\mathcal{D},\mathcal{D}\}} - \pi^{\{\mathcal{C},\mathcal{C},\mathcal{D}\}} > 0$

Proof. Denote by e_h and e_l the efforts under $\mathcal{C}$ and regime $\{\mathcal{C},\mathcal{C},\mathcal{D}\}$, that is when $p_h = p$ and $p_l = 0$. The claim is equivalent to

$$\begin{aligned} p^2(2e_{hh} - 2e_h) + 2p(1-p)(e_{hl} + e_{lh} - e_h - e_l) &> 0 \\ pe_{hh} + (1-p)(e_{hl} + e_{lh}) - (1-p)e_l - e_h &> 0 \\ p\frac{rh}{4} + (1-p)r\frac{h^r l^r (h+l)}{(h^r + l^r)^2} - (1-p)e_l - e_h &> 0 \end{aligned}$$

[38] Here, I do not consider the trivial solutions in $p = 0$ and $p = 1$.

[39] In fact, $p = 1/2$ yields a SE-unbiased contest, and $\pi^{\mathcal{D}-\mathcal{C}} = 0$ by Proposition 9.

[40] In order to sign $\frac{\partial^2 e_h(p)}{\partial p^2}$ one could differentiate (4) and (5) would with respect to p twice. It is clear that it would be an expression in f', f'' and f'''. Conditions on f' and f'' are already imposed by (*Reg*). Appropriate conditions on f''' would drive the sign of $\frac{\partial^2 e_h(p)}{\partial p^2}$.

where I used (6) in the last step. Now, use (8) with $p_h = p$ and $p_l = 0$ to eliminate e_h, and obtain

$$\begin{aligned}
p\frac{rh}{4} + (1-p)r\frac{h^r l^r (h+l)}{(h^r+l^r)^2} - (1-p)e_l - (1-p)\frac{h}{l}e_l - p\frac{rh}{4} &> 0 \\
r\frac{h^r l^r (h+l)}{(h^r+l^r)^2} - \frac{h+l}{l}e_l &> 0 \\
r\frac{h^r l^{r+1}}{(h^r+l^r)^2} &> e_l \\
e_{lh} &> e_l \qquad (37)
\end{aligned}$$

where the last step is implied by (6). Now, e_{lh} is the equilibrium when $p_h = 0$ and $p_l = 0$, and e_l is the equilibrium when $p_h = p$ and $p_l = 0$. Therefore, (37) follows from iv) of Proposition 5. ■

Lemma 16 $\pi^{\{\mathcal{D},\mathcal{C},\mathcal{C}\}} - \pi^{\{\mathcal{D},\mathcal{D},\mathcal{D}\}} > 0$

Proof. Denote by e_h and e_l the efforts under $\mathcal{C}$ and regime $\{\mathcal{D},\mathcal{C},\mathcal{C}\}$, that is when $p_h = 0$ and $p_l = 1-p$. The claim is equivalent to

$$\begin{aligned}
2p(1-p)(e_h + e_l - e_{hl} - e_{lh}) + (1-p)^2(2e_l - 2e_{ll}) &> 0 \\
p(e_h + e_l - e_{hl} - e_{lh}) + (1-p)(e_l - e_{ll}) &> 0 \\
e_l + pe_h - p(e_{hl} + e_{lh}) - (1-p)e_{ll} &> 0 \\
e_l + pe_h - pr\frac{h^r l^r (h+l)}{(h^r+l^r)^2} - (1-p)\frac{rl}{4} &> 0
\end{aligned}$$

where I used (6) in the last step. Now, use (8) with $p_h = 0$ and $p_l = 1-p$ to eliminate e_l, and obtain

$$\begin{aligned}
p\frac{l}{h}e_h + (1-p)\frac{rl}{4} + pe_h - pr\frac{h^r l^r (h+l)}{(h^r+l^r)^2} - (1-p)\frac{rl}{4} &> 0 \\
\frac{h+l}{h}e_h - r\frac{h^r l^r (h+l)}{(h^r+l^r)^2} &> 0 \\
e_h &> r\frac{h^{r+1} l^r}{(h^r+l^r)^2} \\
e_h &> e_{hl} \qquad (38)
\end{aligned}$$

where the last step is implied by (6). Now, e_{hl} is the equilibrium when $p_h = 0$ and $p_l = 0$, and e_h is the equilibrium when $p_h = 0$ and $p_l = 1-p$. Therefore, (38) follows from iii) of Proposition 5. ■

Lemma 17 $\pi^{\{\mathcal{D},\mathcal{C},\mathcal{C}\}} - \pi^{\{\mathcal{C},\mathcal{C},\mathcal{C}\}} > 0$

Proof. Now both regimes include some concealment, and hence there are two efforts under $\mathcal{C}$ for each type according to the regime. Denote by $\hat{e}_h$ and $\hat{e}_l$ the efforts under $\mathcal{C}$ and regime $\{\mathcal{D},\mathcal{C},\mathcal{C}\}$, that is when $p_h = 0$ and $p_l = 1-p$, and by $\bar{e}_h$ and $\bar{e}_l$ the efforts under $\mathcal{C}$ and regime $\{\mathcal{C},\mathcal{C},\mathcal{C}\}$, that is when $p_h = p$ and $p_l = 1-p$. The claim is equivalent to

$$\begin{aligned}
p^2(2e_{hh} - 2\bar{e}_h) + 2p(1-p)(\hat{e}_h + \hat{e}_l - \bar{e}_h - \bar{e}_l) + (1-p)^2(2\hat{e}_l - 2\bar{e}_l) &> 0 \\
p^2\frac{rh}{4} + p(1-p)\hat{e}_h - p\bar{e}_h + (1-p)\hat{e}_l - (1-p)\bar{e}_l &> 0
\end{aligned}$$

where I used (6). Now, use twice (8) with $p_h = 0$ and $p_l = 1-p$ to eliminate $\hat{e}_h$, and with $p_h = p$ and $p_l = 1-p$ to eliminate $\bar{e}_h$, and obtain

$$\begin{aligned}
p^2\frac{rh}{4} + (1-p)\frac{h}{l}\hat{e}_l - (1-p)^2\frac{rh}{4} - (1-p)\frac{h}{l}\bar{e}_l + (1-2p)\frac{rh}{4} + (1-p)\hat{e}_l - (1-p)\bar{e}_l &> 0 \\
(1-p)\frac{h}{l}\hat{e}_l - (1-p)\frac{h}{l}\bar{e}_l + (1-p)\hat{e}_l - (1-p)\bar{e}_l &> 0 \\
\hat{e}_l &> \bar{e}_l
\end{aligned}$$

Now, $\hat{e}_l$ is the equilibrium when $p_h = 0$ and $p_l = 1-p$, and $\bar{e}_l$ is the equilibrium when $p_h = p$ and $p_l = 1-p$. Therefore, $\hat{e}_l > \bar{e}_l$ follows from iv) of Proposition 5. ■

C Appendix C. Extensions

In this appendix I analyze some extensions of the above analysis.

C.1 Continuum of Types

The lack of tractability of private information contests is a well-known issue, and even when the type space is binary as assumed here a closed-form solution exists only in case of a symmetric prior. I have been unable to characterize the link between skewness of prior and disclosure policy when the prior is on a continuum type space, but I show a corroboratory example here. Ewerhart (2010) finds that under a particular continuous distribution of types, the (*Logit-CSF*) contest with $r = 1$ does admit a closed-form solution for equilibrium efforts. I use his results to show that the finding of my paper which links skewness of prior and optimal disclosure policy extends at least to his special case of continuum type space.

In particular, Ewerhart's cumulative distribution function is

$$F(v|\bar{v};\underline{v}) = \frac{\ln\left(\sqrt{\frac{(\bar{v}-\underline{v})^2}{2} + 2\left(\bar{v}+\underline{v}\right)} - \frac{\bar{v}+\underline{v}}{2}\right) - \ln\left(\underline{v}\right)}{\ln\left(\bar{v}\right) - \ln\left(\underline{v}\right)}$$

From (18) in Ewerhart (2010), expected aggregate effort under concealment is

$$\pi^{\mathcal{C}} = \frac{\left(\bar{v}-\underline{v}\right)^2}{\left(\bar{v}+\underline{v}\right)\left[\ln\left(\bar{v}/\underline{v}\right)\right]^2} \tag{39}$$

and expected aggregate effort under disclosure is[41]

$$\pi^{\mathcal{D}} = \int_{\underline{v}}^{\bar{v}} \left[\int_{\underline{v}}^{\bar{v}} \frac{xy}{x+y} f(x)dx\right] f(y)dy \tag{40}$$

where $f(x)$ is the derivative with respect to x of the particular cdf proposed by Ewerhart (2010).

Ewerhart cdf is skewed toward low-types, which corresponds to $0 < p < 0.5$ in my setting of binary type space. According to my Theorem 1, one would expect that $\pi^{\mathcal{C}} \geq \pi^{\mathcal{D}}$, which is however difficult to solve analytically given the functional form of the cdf. Hence, I run numerical simulations. In Figure 5 I fix $\bar{v} = 24$ and let $\underline{v}$ change. On the vertical axis I plot $\pi^{\mathcal{C}} - \pi^{\mathcal{D}}$. It can be seen that it is always positive. Trivially, as $\underline{v} \rightarrow \bar{v} = 24$, $\pi^{\mathcal{C}} \rightarrow \pi^{\mathcal{D}}$.[42] Therefore, the cdf proposed by Ewerhart is skewed toward low-types and the optimal disclosure policy under such cdf is one of full concealment.

The example provided here suggests that the link between skewness of prior and optimal disclosure policy carries over to more general prior distributions of types. Nevertheless, it is obviously not possible to make a general argument out of the analysis of this special case.

[41] The expression $\frac{xy}{x+y}$ comes from the fact that in a public information contest with $f(z) = z$ with valuations x and y, the equilibrium efforts are $\frac{x^2y}{(x+y)^2}$ and $\frac{xy^2}{(x+y)^2}$, which sum to $\frac{xy}{x+y}$.

[42] More numerical simulations are available upon request.

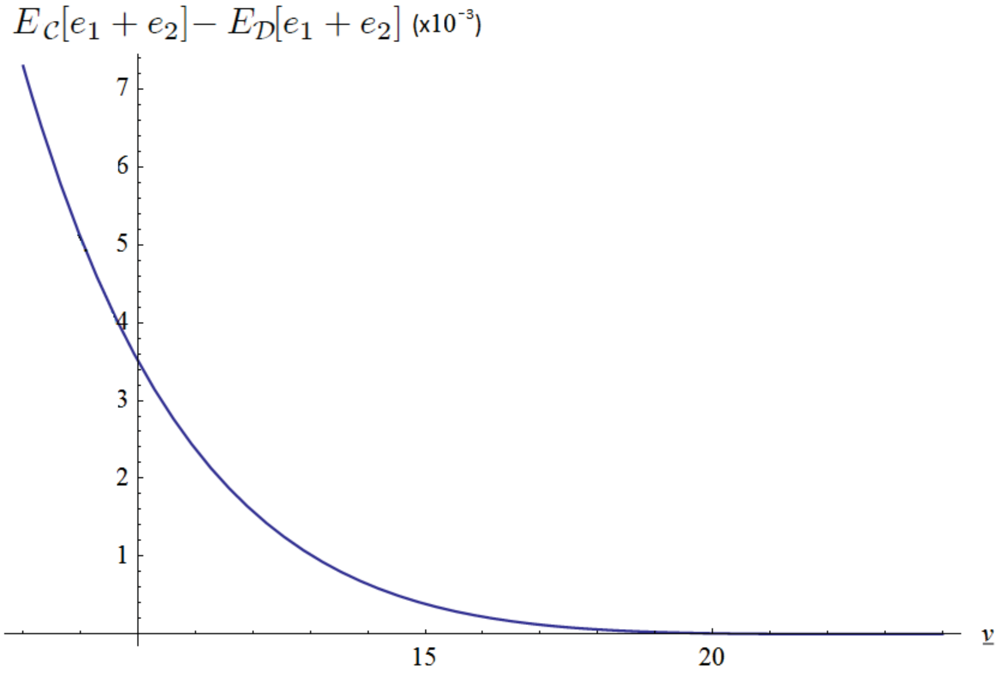

Figure 5:

C.2 Correlated Types

So far types were independent draws from (3). Say there is some positive correlation between types. Then the *principal*'s belief of having a symmetric contest – that is, $\{h, h\}$ or $\{l, l\}$ – increases. Thus, one might conjecture that the principal's incentive to commit to disclose types increases. However, even *contestants*' beliefs are affected by their awareness of a positive correlation: that is, a high-type under positive correlation expects to be facing a high with probability greater than p. Thus, the more types are positively correlated, the less disclosing or concealing $\{h, h\}$ or $\{l, l\}$ affects contestants' efforts. In other words, what matters is how the information the principal discloses is *unexpected* to contestants, rather than *what* is disclosed. In this Section I show that the effects on principal's and contestants' beliefs of correlation balance out in terms of resulting effort, thus leaving the result unchanged.

I define the *correlated-types contest* one where $\Pr(\theta_i = \bar{\theta}|\theta_j = \bar{\theta}) = \alpha \in [0, 1]$ with $j \neq i$ and $\bar{\theta} \in \{h, l\}$. Therefore, $\Pr(\{h, h\}) = p\alpha$, $\Pr(\{l, l\}) = (1 - p)\alpha$, and $\Pr(\{h, l\} \vee \{l, h\}) = 1 - \alpha$. If $\alpha \geq \max\{p, 1 - p\}$, correlation between types is positive (and perfect if $\alpha = 1$). If $\alpha \leq \min\{p, 1 - p\}$, correlation between types is negative (and perfect if $\alpha = 0$). In the remaining cases the sign of correlation depends on type. This special form of symmetric correlation generates what I called in an uncorrelated-setting a SE-unbiased contest, because $p_h = p_l = \alpha$. However, Proposition 9 holds only in our original uncorrelated-setting, and this is why the indifference result under SE-unbiased contest does not hold.[43]

As already mentioned, this correlation is very special, in that it does not allow

[43] Note that the proposed way of modelling type-correlation does not accommodate the case of lack of correlation analyzed in Section 3, unless $p = \alpha = 1/2$.

for no correlation. There exists a closed-form equilibrium effort in $\mathcal{C}$ (see Malueg and Yates, 2004), and it is

$$e_h = rh\left(\frac{\alpha}{4} + (1-\alpha)\frac{l^r h^r}{(h^r + l^r)^2}\right) \tag{41}$$

$$e_l = rl\left(\frac{\alpha}{4} + (1-\alpha)\frac{l^r h^r}{(h^r + l^r)^2}\right) \tag{42}$$

The rest of the proof to find the optimal $\mathcal{P}$ can be easily worked out as a simplified version of the proof of Theorem 1, where the simplification comes from having a closed-form solution for $\mathcal{C}$-equilibrium efforts: (41) and (42). I sketch here the algebra.

$$\pi^{\mathcal{D}-\mathcal{C}} = 0$$

$$2p\alpha\left[e_{hh} - e_h\right] + 2(1-p)\alpha\left[e_{ll} - e_l\right] + (1-\alpha)[e_{hl} + e_{lh} - e_h - e_l] = 0$$

$$p\alpha\tfrac{rh}{4} + (1-p)\alpha\tfrac{rl}{4} + (1-\alpha)\tfrac{r(h+l)h^r l^r}{2(h^r+l^r)^2}$$

$$-r\left[2p\alpha h - 2(1-p)\alpha l - (1-\alpha)h - (1-\alpha)l\right]\left(\tfrac{\alpha}{4} + (1-\alpha)\tfrac{l^r h^r}{(h^r+l^r)^2}\right) = 0$$

using (6), (41) and (42),

$$\alpha(h^r + l^r)^2\left[ph + (1-p)l\right] + 2(1-\alpha)(h+l)h^r l^r$$

$$+\left[h + l + (l-h)\alpha(1-2p)\right]\left(\alpha(h^r+l^r)^2 + 4(1-\alpha)l^r h^r\right) = 0$$

$$\alpha(h^r + l^r)^2\left[ph + (1-p)l + h + l + (l-h)\alpha(1-2p)\right]$$

$$+2(1-\alpha)(h+l)h^r l^r + \left[h + l + (l-h)\alpha(1-2p)\right]\left(4(1-\alpha)l^r h^r\right) = 0$$

$$\alpha\left[h^r + l^r\right]^2\left[(l-h)(1-\alpha)(1-2p)\right] - 4(1-\alpha)h^r l^r(l-h)\alpha(1-2p) = 0$$

whose unique solution is $p = \frac{1}{2}$. thus, I proved that in the correlated-types contest, $\pi^{\mathcal{D}-\mathcal{C}} = 0$ iff $p \in \{0, \frac{1}{2}, 1\}$. This proof mirrors Step 1 of the Proof of Theorem 1. Proofs of Step 2 and 3 are trivial and I omit them.

C.3 Ex-post Disclosure

My findings rely on the principal's ex-ante *commitment* power to disclose types. If the principal lacks such commitment power, but still she can verifiably disclose types, then the disclosure policy is effectively chosen *ex-post*; that is, the disclosure policy if a mapping from type realizations to a disclosure/concealment choice, $\mathcal{P}$: $\{\{h,h\},\{h,l\},\{l,h\},\{l,l\}\} \rightarrow \{\mathcal{D},\mathcal{C}\}^3$. The ex-ante partial information disclosure analyzed in Section 4 has de facto the same mapping, except with commitment, and thus it dominates any ex-post disclosure choice.[44] Not as trivial is the comparison of ex-ante disclosure with only extreme policies (Section 3) and ex-post disclosure: in the former the principal has commitment power but the disclosure policy is rougher – that is, $\mathcal{P} \in \{\mathcal{D}, \mathcal{C}\}$ rather than $\mathcal{P} \in \{\mathcal{D}, \mathcal{C}\}^3$. In this Section I analyze this comparison.

Ex-post, it is a dominant strategy for the principal to disclose if $\{h, h\}$ or $\{l, l\}$, and to conceal otherwise, but ex-post concealment conveys itself information to contestants: a high-type observing ex-post concealment infers of being against a low-type ("if I was against another high-type, the principal would have disclosed"), and viceversa. Hence, contestants perfectly infer types, and thus exert $\mathcal{D}$-equilibrium-efforts. In Section 3 the principal ex-ante commits to either $\mathcal{C}$ or $\mathcal{D}$, and the latter yields the same efforts of the ex-post disclosure choice, as just discussed. Therefore, ex-ante choice of disclosure policy dominates ex-post choice. In other words, if the principal was capable of choosing ex-ante or ex-post the disclosure policy, she would strictly prefer ex-ante decision and commits to $\mathcal{C}$ if $p \in (0, 0.5)$, and she would be indifferent between ex-ante and ex-post for the other values of p.

Nevertheless, as the reader might have noticed, the fact that under ex-post disclosure contestants perfectly infer types is an artifact of the *binary* type space: only in $\{h, l\}$ the principal wants to conceal types.[45] Thus, it would be interesting to check the robustness of this perfect inference result to a type space with 3 or more types, such that contingencies of uneven contest are multiple – $\{h, m\}$, $\{h, l\}$, and $\{m, l\}$ –, and hence the principal might want to conceal in some of them, making inference on the rival's type imperfect. I do not take this avenue here.[46]

C.4 Principal's objective function

Throughout the paper, I assumed that the principal's objective function is the sum of contestants' efforts. This assumption fits situations when the principal equally cares about every contestant's effort, or wants to stimulate the market. Yet, one may think of several situations where the principal is rather concerned with the quality of the winning submission. For instance, in an architectural contest the principal will eventually only implement the winning project, and throw away the losing projects. Using the notation of the present paper, the suitable objective function in such situations would be the expected winning effort (shortly, EWE): $p_1(e_1, e_2) * e_1 + p_2(e_2, e_1) * e_2$. Optimal contest design under EWE-maximization is proposed and studied by Serena (2015). How does the optimal disclosure policy change, if any?

Consider a (*Logit-CSF*) contest with $r = 1$, and consider 2 situations: (A) where $e_1 = e_2 = 5$, and (B) where $e_1 = 7$ and $e_2 = 2$. A principal maximizing

[44] Beliefs and efforts inducible by means of an ex-post disclosure policy can be induced by means of a partial information disclosure with commitment, whereas the opposite is not true.

[45] Note that contingencies $\{h, l\}$ and $\{l, h\}$ are equivalent.

[46] When a sender chooses a verifiable message to send to a receiver and the sender has monotone preferences in the receiver's action, then it is well-known that perfect inference occurs in every sequential equilibrium, see Milgrom (1981) for the first of these types of results. I am not aware of an extention of this classic unravelling result which accommodates my two-receiver setting. See Section II.1 in Dranove and Jin (2010) for a survey on verifiable message disclosure games.

aggregate effort would go for (A) because $5+5>7+2$. A principal maximizing EWE would go for (B) because $\frac{5}{5+5}5+\frac{5}{5+5}5=5<\frac{7}{7+2}7+\frac{2}{7+2}2=\frac{53}{9}$. The sole goal of this numerical example is to convince that EWE-maximization makes the principal favour heterogeneity of efforts. Now, situations of heterogeneity of efforts arise in my setting when contestants happen to be of asymmetric types, $\{h,l\}$ or $\{l,h\}$, and in those situations the principal ideally wants to have concealed types. Therefore, EWE-maximization increases the incentive to conceal types as opposed to the standard aggregate effort maximization. In fact, it could be proven that the value of p which makes the principal indifferent between $\mathcal{D}$ and $\mathcal{C}$ of Theorem 1 is strictly greater than $p=\frac{1}{2}$.[47]

I acknowledge at least other two sensible alternative objectives. First, if one has in mind harmful conflicts or wasteful lobbying, the principal would probably want to minimize aggregate efforts. In this case, the results retrieved in this paper are simply inverted (for instance, turn Figure 2 upside down). Second, if one has in mind contests for hiring new employees or the economic job market, and in general a contest aiming at selection of the most skilled contestant as the winner, then a suitable principal's objective would be to maximize the probability of a high-type winner. This last case, although more convoluted, yields the same optimal disclosure policy of the aggregate effort minimization case – that is, $\mathcal{D} \iff p\in(0,\frac{1}{2})$, $\mathcal{C} \iff p\in(0,\frac{1}{2})$, and indifference in the remaining cases.[48]

C.5 Beyond two players

Having 3 or more players simultaneously in the contest creates technical problems even if focusing only on interior equilibrium and leaving the type-space binary. The reason is that the more players we add, the more we lose the clear-cut monotonicities which allowed us to analytically characterize the results. This can be already seen in a simple 3-player case.

For instance, in a (*logit-CSF*) contest with $r=1$, $l=1$, and $h=2$, one the one hand it holds that $e_{lll}>e_{llh}>e_{lhh}$ as expected,[49] but on the other hand it also holds that $e_{hhl}>e_{hll}>e_{hhh}$.[50] This implies that symmetry of types does not necessarily

[47] Thus, the theorem would read as follows. Under (*Logit-CSF*), possibility of full-disclosure or full-concealment only, and expected winning effort maximization, the optimal disclosure policy for the principal is to commit to full concealment if $p\in[0,\bar{p}]\cup\{1\}$ where $\exists!\bar{p}\in(\frac{1}{2},1]$, and to commit to full disclosure otherwise.

I omit the proof because long and convoluted. My efforts to single out a closed form solution for $\bar{p}$ were not successful. However, numerical simulations show that for most parameter values $\bar{p}=1$. That is, full concealment is optimal $\forall p$.

[48] An easy way yo see this is as follows. The principal's disclosure policy affects the probability of a high type winner only in asymmetric contingencies - i.e., $\{h,l\}$ and $\{l,h\}$. Asymmetric contingencies has a probability of occurrence equal to $2p(1-p)$, which equals 0 when $p\in\{0,1\}$ and is strictly positive when $p\in(0,1)$. Thus, besides the usual indifference when $p\in\{0,1\}$, full concealment dominates full disclosure if and only if $\frac{e_h}{e_h+e_l}\geq\frac{e_{hl}}{e_{hl}+e_{lh}}=\frac{h}{h+l}$, or equivalently $\frac{e_h}{e_l}\geq\frac{h}{l}$. Since e_h is increasing in p and e_l is decreasing in p (see Figure 1), the ratio $\frac{e_h}{e_l}$ is increasing in p. By Corollary 8, $\frac{e_h}{e_l}$ equals $\frac{h}{l}$ exactly when $p=\frac{1}{2}$. By these last two facts, $\frac{e_h}{e_l}\gtreqless\frac{h}{l} \iff p\gtreqless\frac{1}{2}$, and the result follows.

[49] Approximating to the second digit, $e_{lll}=0.20, e_{llh}=0.16, e_{lhh}=0$

[50] Approximating to the second digit, $e_{hhl}=0.5, e_{hll}=0.48, e_{hhh}=0.44$

imply greater efforts. The reason for this new ranking is that the set of rival's types that maximizes effort of a high-type in the trade-off between having too weak or too strong rivals is a high and a low. Whereas, having two low rivals implies less competitive of a battlefield, and having two high-type rivals implies too competitive of a battlefield. These ranking depends on the parameters of the model (r, l, h), on the contrary of the findings of the 2-player setting (Theorem 1 and Theorem 2). Given that evenness of playing-field does not necessarily yield more effort, the signs of SE are also ambiguous. In the above parametrical example, on the one hand the +SE in e_h is standard, since it is the best reply to standardly ranked $\mathcal{D}$-efforts (that is, $\lim_{p\to 0} e_h = 0.498 > e_{hll}$), on the other hand the SE in e_l is unexpected, since it is the best reply to non-standardly ranked $\mathcal{D}$-efforts (that is, $\lim_{p\to 1} e_l = 0.05 < e_{lhh}$). In other words, if $p \to 0$, a high-type knows of being against two low either aware of being in a $\{h, l, l\}$-contest under $\mathcal{D}$, or believing of being in a $\{l, l, l\}$-contest under $\mathcal{C}$. Thus, this high-type best replies to e_{llh} under $\mathcal{D}$, or best replies to e_{lll} under $\mathcal{C}$. Because $e_{llh} < e_{lll}$, $e_h > e_{hll}$ as expected. Alternatively, if $p \to 1$, a low-type knows of being against two high either aware of being in a $\{l, h, h\}$-contest under $\mathcal{D}$, or believing of being in a $\{h, h, h\}$-contest under $\mathcal{C}$. Thus, this high-type best replies to e_{hhl} under $\mathcal{D}$, or best replies to e_{hhh} under $\mathcal{C}$. Because $e_{hhl} > e_{hhh}$, $e_l > e_{lhh}$ as unexpected.

References

- Aoyagi, M., 2010. "Information feedback in a dynamic tournament," *Games and Economic Behavior*, Elsevier, vol. 70(2), 242-260.
- Baik, K., 1994. “Effort Levels in Contests with Two Asymmetric Players,” *Southern Economic Journal*, 61, 367–378.
- Brown, J., 2011. "Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars", *Journal of Political Economy.* 119(5): 982-1013
- Calsamiglia, C., Franke, J., and Rey-Biel, P. 2013. "The Incentive Effects of Affirmative Action in a Real Effort Tournament", *Journal of Public Economics*, 98, 15-31
- Corchón, L., 2000. "On the allocative effects of rent-seeking", *Journal of Public Economic Theory*, vol. 2(4), 483-491.
- Denter, P., Morgan, J., and Sisak, D., 2011. "Where Ignorance is Bliss, 'tis Folly to be Wise: Transparency in Contests", Working Paper.
- Dixit, A., 1987. "Strategic Behavior in Contests", *The American Economic Review*, vol. 77 (5), 891-898
- Dranove, D., and Jin, G. Z., 2010. "Quality Disclosure and Certification: Theory and Practice", *Journal of Economic Literature*, Vol. 48, No. 4, 935-963

- Dubey, P, 2013. "The role of information in contests", *Economics Letters*, vol. 120(2), 160-163.

- Einy, E., Haimanko, O., Moreno, D., Sela, A., and Shitovitz, B., 2016. " Equilibrium Existence in Tullock Contests with Incomplete Information", *Journal of Mathematical Economics*, forthcoming.

- Ewerhart, C., 2010. "Rent-seeking contests with independent private values", Working Paper.

- Ewerhart, C., and Quartieri, F., 2013. "Unique Equilibrium in Contests with Incomplete Information", Working Paper.

- Fu, Q., Jiao, Q., and Lu, J., 2011. "On disclosure policy in contests with stochastic entry", *Public Choice*, vol. 148(3), 419-434.

- Fu, Q., Jiao, Q., and Lu, J., 2014. "Disclosure policy in a multi-prize all-pay auction with stochastic abilities", *Economics Letters*, vol. 125(3), 376-380.

- Fullerton, R. L. and McAfee, R. P. 1999. "Auctioning Entry into Tournaments", *The Journal of Political Economy*, vol. 107, Issue 3, 573-605.

- Hurley, T., and Shogren, J., 1998a. "Asymmetric information contests", *European Journal of Political Economy*, vol. 14(4), 645–665.

- Hurely, T., and Shogren, J. 1998b. "Effort levels in a Cournot Nash contest with asymmetric information", *Journal of Public Economics*, 69, 195–210.

- Kamenica, E., and Gentzkow, M. 2011. "Bayesian Persuasion", *American Economic Review*, 101(6): 2590-2615.

- Kovenock, D., Morath, F., and Münster, J. 2015. "Information sharing in contests", *Journal of Economics & Management Strategy*, 24, 570 -596.

- Levitt, S. D., 1994. "Using Repeat Challengers to Estimate the Effect of Campaign Spending on Election Outcomes in the U.S. House", *Journal of Political Economy*, 102(4), 777-98

- Lim, W. and Matros, A., 2009. "Contests with a stochastic number of players", *Games and Economic Behavior*, vol. 67(2), 584-597.

- Malueg, D., and Yates, A., 2004. "Rent seeking with private values", *Public Choice*, vol. 119, 161–178.

- Milgrom, P., 1981. "Good news and bad news: representation theorems and applications", *Bell Journal of Economics*, vol. 12(2), 380–391.

- Milgrom, P., and Weber, R., 1982. "A Theory of Auctions and Competitive Bidding", *Econometrica*, vol. 50(5), 1089-1122.

- Milkovich, G. T., and Newman, J. M., 1996. "Compensation". Chicago: McGraw-Hill/Irwin

- Morath, F. and Münster, J., 2013. "Information acquisition in conflicts", *Economic Theory*, 54, 99-129.
- Morgan, J., and Várdy, F., 2007. "The value of commitment in contests and tournaments when observation is costly", *Games and Economic Behavior,* vol. 60, 326–338.
- Myerson, R., and Wärneryd, K., 2006. "Population uncertainty in contests", *Economic Theory*, vol. 27(2), 469-474.
- Nti, K. O., 1999. "Rent-seeking with asymmetric valuations", *Public Choice*, vol. 98(3-4), 415-430.
- Serena, M., 2015. "Quality contests", Working Paper.
- Skaperdas, S., 1996. "Contest Success Functions", *Economic Theory*, vol. 7(2), 283-290.
- Suijsa, J., and Wielhouwerb, J. L., 2012. "Disclosure Regulation in Duopoly Markets: Proprietary Costs and Social Welfare", *European Accounting Review*,Volume 23, Issue 2, 227-255.
- Szech, N., 2011. "Asymmetric All-Pay Auctions with Two Types", Working Paper.
- Tullock, G., 1980. "Efficient Rent Seeking", in J. Buchanan, R. Tollison, and G. Tullock, *Towards a Theory of a Rent-Seeking Society*, College Station: Texas A&M University Press.
- Warneryd, K., 2003. "Information in Conflicts", *Journal of Economic Theory*, Volume 110, 121-136
- Wasser, C., 2013. "Incomplete information in rent-seeking contests", *Economic Theory*, Volume 53, Issue 1, 239-268
- Yan, J., and Yin, X., 2012. "Information Acquisition and Disclosure: the Case of Differentiated Good Duopoly", Working Paper.
- Yildirim, H., 2005. "Contests with Multiple Rounds", *Games and Economic Behavior,* vol. 51(1), 213–227.
- Zhang, J. and Zhou, J. "Information Disclosure in Contests: A Bayesian Persuasion Approach", *Economic Journal*, forthcoming.

Chapter 2
Quality Contests

1 Introduction

What are plausible objectives of a contest designer? First, and most commonly adopted in the literature, is the sum of contestants' efforts. This objective fits well situations where the designer aims to stimulate overall activity in a specific field. A CEO designs a job promotion contest in order to increase the firm's profit. A ministry of education organizes mathematical olympiads in order to encourage national scientific aptitude. Second, is competitive balance, understood as the uncertainty of the contest outcome. The organizer of a sport event benefits from outcome uncertainty in that it thrills the interest of the audience, and thus it increases the organizer's sales. Third, is the need to attract participation - as in online communities where users contribute to the content - or deter participation - as in conflicts.[1] Fourth, is the quality of the winner's entry. This is suitable in contests where each contestant submits an entry, but only the winner's entry is eventually implemented by the designer. A mayor runs an architectural contest to build a new bridge in town, and two architects submit their projects.[2] Each architect's project is of a certain endogenous and costly *quality*. The mayor will eventually have only one of the two projects built. If the mayor perfectly observes the quality of the projects (i.e., perfectly discriminatory contest, or all-pay auction), the suitable objective of the mayor is to maximize the highest quality between the two projects, because the project of highest quality wins with certainty and it is eventually built. If instead the mayor observes the quality of the projects with some noise (i.e., imperfectly discriminatory contest),[3] then there is some positive probability that the mayor makes a mistake and have the low quality bridge built. In this realm of noisy contests, it seems reasonable to say that the mayor maximizes the quality of the *winning* bridge, which is the one eventually built, rather than the quality of the *highest quality* bridge, because she could be making a mistake when evaluating entries due to the noise. Thus, we claim that, in noisy contests, the expected winning quality, is the natural objective of a designer who benefits only from the winning entry.[4] Throughout the paper we conform to the literature and use the conventional word "effort" to define the costly and sunk investment of contestants, despite words quality, bid and effort are interchangeable. We compare the optimal contest design under expected winning

[1] A comprehensive literature review of proposed objective functions is beyond the scope of this work. Papers focusing on the sum of efforts are, for instance, Baye, Kovenock, and de Vries (1993), Moldovanu and Sela (2001), Moldovanu, Sela and Shi (2007), and Brown (2011). As for competitive balance, see Falconieri (2004), Palomino (2004), and Vrooman (2012). Contests targeting participation are, for instance, in Azmat and Möller (2009).

[2] Other applications of contests where the designer is likely to benefit from the best entry are product development race, research contest, or even the job market for economists, where the recruiting department wants to hire a candidate whose entry (the job market paper) is ideally the best one, because the candidate's entry is likely to be correlated with the candidate's future output.

[3] The noise could arise for instance when the time spent analyzing the details of the two projects in order to select the winner is limited or costly.

[4] This idea mirrors the highest bid maximization in all-pay auctions: see, for instance, Cohen et al. (2008), Segev and Sela (2014), or Jonsson and Schmutzler (2015). Maximization of highest bid in noisy contest would correspond to maximize the quality of the best project, but the best project could lose because of the noise, and if this is the case, we claim its quality is of no value to the designer. Expected winning quality takes this into account.

effort maximization and under the conventional sum of effort maximization.[5]

In order to easily recall throughout the paper these two different objectives, we name *stimulative contest* the conventional setting in which the designer maximizes the sum of contestants' efforts (thus aiming at stimulating competition), and *quality contest* the newly proposed setting in which the designer maximizes the expected winning effort. In the first period, the designer chooses and commits to a set of "rules" which shapes the competitive environment that contestants face. In the second period, the noisy contest is played. The way we add noise to the winner selection process is by means of a standard lottery contest success function - see (1) - which is a tractable reduced-form corresponding to an inverse exponential distribution of noise, see Jia (2008). We cherry-pick three rules of contest design for which the different results on how to optimally design a stimulative or a quality contest most greatly help grasp the intuitions behind the differences between these two types of contest. Retrieving these results and delivering these intuitions is the main goal of the paper.

First, we analyze whether and how the designer wants to *level the playing field* by giving, for instance, a competitive advantage to the "low type" - where type is marginal cost of effort. We find that, on the contrary of a stimulative contest, in a quality contest it is optimal to leave the playing field unlevelled, to some extent. The intuition is as follows. Start from a contest between two contestants - a high type and a low type - with perfectly levelled playing field, such that both contestants have the same probability of winning in equilibrium and competition for the prize is maximum. Let us move away from the perfectly levelled playing field, such that one of the two contestants is now more likely to win in equilibrium. Then, two effects simultaneously arise: (i) both contestants' efforts decrease because competition for the prize is no longer maximum, and (ii) the probability of winning of one contestant increases, and that of the other decreases. A stimulative contest is harmed by (i) and not affected by (ii), and hence an unlevelled playing field is always detrimental in stimulative contests, as found in the existing literature - see Nti (2004), Runkel (2006), Franke (2012) and Franke et al (2013). On the contrary, a quality contest is harmed by (i), but benefits from (ii) if the high type is the one given the advantage, because this increases the probability that the high type's effort (which is greater than the one of the low type) is the winning one. This trade off present only in quality contests explains the optimality of leaving some degree of advantage to the high type (i.e., non-perfectly levelled playing field).

Second, we study the *optimal selection of contestants*.[6] The designer observes a set of contestants - "applicants" - of known types, among which she chooses the contestants - "finalists" - who are granted access to the contest. Once the choice of finalists is made and publicly announced, some of them might still drop out of

[5] To the best of our knowledge, there are no works performing such analysis. Yet, we acknowledge that, at the end of page 345 of Andreff and Szymanski (2006), the authors mention "[...] there are also situations in sports where the aim of the contest designer is not to maximise total expected effort but expected effort of the winner". However, they do not run any analysis about it. Drugov and Ryvkin (2016) propose as the principal's objective function the expectation over types of a general function that may depend on efforts, types, and bias of the playing field. Expected winning effort is thus a special case.

[6] All contest design rules are considered separately. Thus, for instance, when the designer selects contestants, she cannot simultaneously level the playing field.

the contest if better off doing so. That is, if the other selected finalists are of much greater type, a finalist is better-off quitting. In a quality contest, we find that it may be profitable to exclude some contestants, especially when types are sufficiently homogenous. In fact, types homogeneity across contestants yields efforts homogeneity across contestants, and when efforts are homogeneous, expected winning effort equals individual effort regardless of who wins. Individual effort decreases in the number of finalists, because each finalist has roughly $1/n$ probability of winning the prize, and thus as n increases, individual willingness to exert effort decreases. Thus, if contestants are sufficiently homogenous, excluding some applicants from the final is profitable in a quality contest.[7] However, the sum of efforts - on the contrary of the expected winning effort - decreases as contestants quit. This negative effect turns out to be stronger than any positive effect on individual efforts that exclusions may generate, and this is why Fang (2002) finds that no exclusion is profitable in stimulative contests, regardless of the heterogeneity of types. Thus, while the designer of a stimulative contest should not be concerned about excluding contestants, and rather be concerned about stimulating applications, we find that the designer of a quality contest might want to exclude some applicants from the contest.[8]

Third, since we found that not only exclusion is often profitable in quality contests, but for a large set of contestants' types the optimal set of finalists is two, it seems natural to run *comparative statics on contestants' types* in a two-contestant contest. We find that while weakening the favorite - i.e., an increase in her marginal cost of effort - could be beneficial both in a quality contest and in a stimulative contest, weakening the underdog could be beneficial only in a quality contest, despite it lowers both contestants' efforts. The intuition builds upon the above-mentioned playing field effect, and adds the layer of individual effect; that is, changing a contestant's type affects the playing field - making it more or less levelled -, as well as the affordability of effort of that contestant (individual effect). In particular, weakening the underdog is beneficial to a quality contest when types are sufficiently heterogenous and the playing field is sufficiently levelled. The intuition of this result is given in Section 6.

The structure of the paper is as follows. In Section 2 we spell out the model and formalize the difference between a stimulative and a quality contest. In Section 3 we propose a way to disentangle in equilibrium efforts the effect due to the contestant's individual type and the effect due to the contest playing field. In sections 4, 5 and 6 we build upon Section 3 to analyze the optimal contest design's differences between a stimulative and a quality contest, in particular: optimal levelling of playing field (Section 4), optimal contestants' selection (Section 5), and comparative statics on contestants' types (Section 6). Section 7 concludes. Proofs are in the Appendix.

[7] If all contestants are *perfectly* homogenous, the optimal number of contestants is 2.

[8] Our model and Fang (2002) adopt a lottery contest success funtion. If we move away from the realm of lottery contests and enter into all-pay auctions, our finding of profitability of exclusion "from the bottom" - that is, starting from the weakest contestants - is in stark contrast with the exclusion principle proposed by Baye et al (1993). They find that in an all-pay auction excluding top contestants might be beneficial to boost competition among weaker contestants. However, their objective function is the sum of efforts. In Appendix B we compare all-pay auctions and noisy contest under expected winning effort maximization.

2 Model

Consider a complete information contest with n risk-neutral contestants indexed by $i \in \{1, ..., n\}$. Contestants compete for a prize whose value we normalize wlog to 1. Each contestant i simultaneously chooses effort level e_i, and has a probability of winning the prize equal to

$$p_i(e_1, ..., e_n) = \begin{cases} 0 & \text{if } e_j = 0 \ \forall j \in \{1, ..., n\} \\ \frac{\alpha_i e_i}{\sum_{j=1}^{n} \alpha_j e_j} & \text{otherwise} \end{cases} \tag{1}$$

with $p_i : \mathbb{R}^n_+ \to [0, 1]$, and $\alpha_i \geq 0 \ \forall i \in \{1, ..., n\}$. The cost of effort is linear, and the possibly heterogeneous marginal cost is referred to as contestant's type. Hence, contestant i chooses e_i to maximize

$$\frac{\alpha_i e_i}{\sum_{j=1}^{n} \alpha_j e_j} - \beta_i e_i \tag{2}$$

Prior to contestants' simultaneous choice of efforts, the contest designer commits to contest rules $\mathcal{R}$. In Section 4, the rule $\mathcal{R}$ to be chosen is (α_1, α_2), with $n = 2$; that is, the designer gives a competitive advantage (or disadvantage) to contestant i, or the designer "levels" (or "unlevels") the playing field. In Section 5, the rule $\mathcal{R}$ to be chosen is $\alpha_i \in \{0, 1\}$ for $n \geq 2$; that is, the designer selects the set of contestants who can participate in the contest, but cannot level the playing field.[9] In Section 6, we impose $n = 2$, and run comparative statics on β_i's.

As discussed, our focus is on the objective function of the designer, which is either the sum of efforts, denoted by $\widehat{sum}$, as prevalent in the literature

$$[\textbf{stimulative contest}] \quad \widehat{sum} = \sum_{i=1}^{n} e_i \tag{3}$$

or the expected winning effort, denoted by $\widehat{ewe}$:

$$[\textbf{quality contest}] \quad \widehat{ewe} = \sum_{i=1}^{n} p_i(e_1, ..., e_n) e_i \tag{4}$$

In order to easily recall along the paper these two alternative objective functions, we name the contest accordingly. In a stimulative contest the designer maximizes (3), and in a quality contest the designer maximizes (4).

Prior to discussing optimal $\mathcal{R}$, we analyze in the next section equilibrium efforts, and we propose a notation which is convenient to deliver intuitions on the optimal $\mathcal{R}$ later on.

[9] Note that if $\alpha_i = 0$, then $e_i^* = 0$, and thus choosing $\alpha_i = 0$ is equivalent to excluding contestant i from the contest.

3 Equilibrium with $n = 2$

The unique NE of the contest described above with 2 contestants is well-known in the literature[10]

$$e_i^* = \frac{\alpha_1\alpha_2\beta_1\beta_2}{(\alpha_1\beta_2 + \alpha_2\beta_1)^2} * \frac{1}{\beta_i}$$

We define the following

$$\begin{aligned} \alpha &= \frac{\alpha_1}{\alpha_2} \\ \beta &= \frac{\beta_1}{\beta_2} \\ S &= \frac{\alpha}{\beta} \end{aligned}$$

So that equilibrium efforts can be written as

$$e_i^* = \underbrace{\frac{S}{(1+S)^2}}_{\text{playing field effect}} * \underbrace{\frac{1}{\beta_i}}_{\text{individual effect}} \tag{5}$$

and the equilibrium probabilities of winning as

$$p_1^* = \frac{S}{S+1} \quad \text{and} \quad p_2^* = \frac{1}{S+1} \tag{6}$$

Thus, we name contestants according to (6): we name *the favorite* the contestant more likely to win in equilibrium, and we name *the underdog* the contestant less likely to win. The value of S determines who is the favorite among contestants (i.e., $S \gtreqless 1 \Leftrightarrow \alpha \gtreqless \beta \Leftrightarrow p_1^* \gtreqless p_2^*$). We refer to the contestant with the lowest (greatest) β_i as the high (low) type. Thus, being a high or low type is different from being the favorite or the underdog, and for instance contestant 1 might be the high type (when $\beta \leq 1$), and yet less likely to win in equilibrium than her rival if the effect of α over-compensate the asymmetry of types (when $\alpha \leq \beta$).

This notation allows to distinguish in the equilibrium efforts (5) between:

- the playing field effect: $\frac{S}{(1+S)^2}$. This term is maximized at $S = 1$, that is when $p_1^* = p_2^* = \frac{1}{2}$. This situation is often referred to as "perfectly levelled playing field".
- the individual effect: $\frac{1}{\beta_i}$. This term accounts for the contestant's idiosyncratic type: the greater is a contestant's marginal cost of effort, the more severe is the negative individual effect.

Changes in α_i's (Section 4) have only playing field effects - through S. Changes in β_i's (Section 6) have both playing field effects - through S - and individual effects.

Substitute (5) into (3) in order to write the sum of efforts as

$$\widehat{sum} = \frac{S}{(1+S)^2}\left(\frac{1}{\beta_1} + \frac{1}{\beta_2}\right) \tag{7}$$

[10] See for example Nti (2004).

and substitute (5) into (4) in order to write the expected winning effort as

$$\widehat{ewe} = \frac{S}{(1+S)^3}\left(\frac{S}{\beta_1} + \frac{1}{\beta_2}\right) \tag{8}$$

Note that in (8) the playing field effect and individual effect are not (multiplicatively) separable. That is, in quality contests playing field and individual effects are intertwined, as it will be clear in the analysis below.

With the preliminaries developed in this section in our hands we analyze the optimal $\mathcal{R}$ in the next sections.

4 Levelling the playing field

In this section the designer can give a competitive advantage (or disadvantage) to one of the two contestants by choosing the parameter $\alpha \geq 0$.[11] There is abundance of examples of such policies: court-appointed attorney in trials where a party cannot afford a private attorney, governmental subsidies to start-ups, public procurement auctions with bidding advantage depending on the firm size, sport competitions where the weak - or losing - player is given a competitive advantage,[12] hiring procedures of international organizations that guarantee a minimal share of employees from developing countries, and so on.

In a stimulative contest, the problem $\max\limits_{\alpha} \widehat{sum}$ is equivalent to $\max\limits_{\alpha} \frac{S}{(1+S)^2}$ because α enters $\widehat{sum}$ only in S. The maximum of $\frac{S}{(1+S)^2}$ is in $S = 1$, or equivalently when $\alpha = \beta$. This result is already present, for instance, in Nti (2004).

Proposition 1 *[Nti (2004)] The optimal α in a stimulative contest is*

$$\alpha^p = \beta$$

Intuitively, the optimal design of a stimulative contest is when competition between contestants is maximum, and this is achieved by setting an α such that contestants are equally likely to win in equilibrium ($S = 1$) regardless of their marginal costs. We denote by α^p the optimal α to refer to the *perfect* levelling of the playing field.

In a quality contest, we retrieve the optimal α building upon the analysis of Section 3.

Proposition 2 *The optimal α in a quality contest is*

$$\alpha^q = \beta\left(1 - \beta + \sqrt{(1-\beta)^2 + \beta}\right) \tag{9}$$

and thus, $\beta \lesseqqgtr 1 \iff \alpha^q \gtreqqless \alpha^p = \beta \iff p_1^* \gtreqqless p_2^*$.

[11] Note that one of the α_i's can be normalized to 1, and in the special case of $n = 2$ the choice of (α_1, α_2) coincides with the choice of α.

[12] Example of these are golf tournaments where the performance of low-performing players is normalized, or horse handicap race where horses carry different weights and a faster horse is given a heavier weight, or football where the ball is given to the team who just conceded a goal.

Thus, in an optimally levelled quality contest the high type is more likely to win in equilibrium than the low type. The intuition is as follows. Each contestant's effort is maximum when the playing field is perfectly levelled, that is $\alpha^p = \beta$. Setting α lower or greater than β would decrease both e_1 and e_2 (i.e., negative playing field effect). This would decrease the sum of efforts $e_1 + e_2$, and thus moving away from $\alpha^p = \beta$ in a stimulative contest is harmful (Proposition 1). On the other hand, the objective function of a two-contestant quality contest is $p_1e_1 + p_2e_2$,[13] thus it is not trivial that setting α lower or greater than β - which decreases both e_1 and e_2 - is detrimental, because p_1 and p_2 also change. Say wlog that $\beta \leq 1$, such that 1 is the high type, 2 is the low type, and $e_1 \geq e_2$ (see (5)). Setting α lower than β is not optimal because, besides the negative playing field effect of decreasing e_1 and e_2, it also decreases p_1 and increases p_2 (in fact, $\alpha < \beta \implies S < 1 \implies p_1^* < p_2^*$, see (6)) which has a negative effect on $p_1e_1 + p_2e_2$ because $e_1 \geq e_2$. Setting α greater than β would instead increase p_1 and decrease p_2, which increases $p_1e_1 + p_2e_2$ because $e_1 \geq e_2$. Thus, while the detrimental decrease of e_1 and e_2 occurs both as α becomes greater or lower than β, the change of p_1 and p_2 are detrimental if α decreases from β and beneficial if α increases from β. For sufficiently small increase of α from β, the beneficial effect of the changes of p_1 and p_2 dominates the negative playing field effect, and thus $\alpha^q \geq \beta$. The same argument obviously holds if we swap identities: that is, if $\beta \geq 1$, $\alpha^q \leq \beta$. Hence, a quality contest designer levels the playing field, but not perfectly, and leaves the high type more likely to win in equilibrium (Proposition 2).[14]

The following simple corollary of Proposition 2 set boundaries on how unlevelled the playing field could be in an optimally levelled quality contest.

Corollary 3 *In a quality contest,* $\alpha^q \in [\frac{1}{2}\beta, 2\beta]$, $S \in [\frac{1}{2}, 2]$ *and* $p_1^*, p_2^* \in [\frac{1}{3}, \frac{2}{3}]$.

Thus, in an optimally levelled quality contest the playing field is never left too unlevelled. The reason is that a very unlevelled playing field would entail a very acute drop of individual efforts, which cannot be caught up by any possibly positive effect due to changes in p_1 and p_2.[15]

5 Contestants selection

In this section we analyze the optimal selection of contestants by the designer given a set of contestants of known types. No levelling of playing field is now possible. We analyze whether it is optimal to exclude some contestants from the contest.

[13]We often omit dependency of p_i's on e_i's along the paper.

[14]For instance, if $\beta = 2$, $\alpha^q = 2\left(\sqrt{3} - 1\right) \approx 1.46$, thus $p_1^* \approx 0.42$, $p_2^* \approx 0.58$ and $\frac{p_2^*}{p_1^*} = \frac{1}{\sqrt{3}-1} \approx 1.36$. That is, if the high type is twice as strong as the low type ($\beta = 2$), then the optimally levelled quality contest leaves the high type 36% more likely to win than the low type in equilibrium.

[15]One might wonder why, if in general not levelling perfectly the playing field is optimal, distorting an already levelled field is not optimal (i.e., why if $\beta = 1$, $\alpha^q = 1$). When $\beta = 1$, setting $\alpha = 1$ makes contestants exert the *same* effort, say $\bar{e}$, and hence the expected winning effort is $\bar{e}$. Setting an $\alpha \neq 1$ rather than $\alpha = 1$ would make both contestants exert less than $\bar{e}$, and hence the winning effort will certainly be lower than $\bar{e}$ regardless of how p_1 and p_2 change. We thank Mikhail Drugov to raise this point.

The Olympic Committee announces the set of finalists competing for the prize of hosting the Olympic Games. Conference organizers announce the set of finalists out of which the winner of the "Best Paper Award" will be selected. Narrowing down the list of finalists is an especially crucial ingredient in research contests. The Virgin Earth Challenge is a $25 million prize contest to be awarded to whoever can provide sustainable ways of removing greenhouse gases from the atmosphere. On November 2011 the list of 11 finalists among the more than 2600 applications was announced. The Qualcomm Tricorder X PRIZE is a $10 million prize contest to be awarded to whoever can design a portable health diagnostic device of less than 2.3 kg, able to accurately diagnose 12 diseases. On August 2014 the list of 10 finalists was announced. The winners of these two research contests are still to be decided.

An important remark has to be done. Once the choice of contestants is made and announced, contestants might drop out of the contest by exerting zero effort if they find it profitable - namely, if too weak as opposed to the other selected contestants. When selecting contestants, the designer takes into consideration this endogenous participation and compares her payoff - either $\widehat{sum}$ or $\widehat{ewe}$ - under different subsets of contestants.[16]

We first deliver the main intuition of this section with a three contestants example, which is the simplest non-trivial case. Consider a contest between 3 contestants with marginal costs equal to $\beta_1 \geq \beta_2 \geq \beta_3 > 0$. Normalize the marginal cost of the high type, $\beta_3 = 1$. The results of the 3-player optimal contestant selection in a quality contest are in Figure 1.

First, consider contestants' participation choice. If two contestants are selected, no contestant exerts 0 effort in equilibrium. If all three contestants are selected, then the low type - contestant 1 - is better off quitting when she is too weak compared to other two contestants, the medium and the high type. This occurs when β_1 is sufficiently large, in the yellow region below the dotted line. Thus, only medium and high types {2,3} exert positive effort, and the contest is a two-player one.[17] In the remaining region - i.e., between the two parallel lines - the low type is sufficiently strong to be willing to participate, if selected by the designer. This region is in turn separated in a red and a yellow region (above the dotted line).

In the red region, β_1 and β_2 are between 3 and 5, and sufficiently close to each other, whereas $\beta_3 = 1$. That is, the low and medium types are sufficiently similar in types, but significantly weaker than the high type (the "superstar"). Here it turns out to be optimal to admit all the contestants to the contest in order to create competition among the two (similar) weak contestants, which has positive spillovers on the effort of the superstar. Thus, no exclusion is optimal for the designer.

In the yellow region (above the dotted line), the three contestants are sufficiently homogenous in types - that is, β_1 and β_2 are sufficiently close to $\beta_3 = 1$. This means that exerted efforts are also sufficiently similar. Hence, $\widehat{ewe} \approx e_i$ regardless of the number of active contestants - i.e., it matters little who wins because individual efforts are sufficiently similar. In this situation, excluding a contestant yields higher competition among the remaining two, because their individual equilibrium prob-

[16]The analysis in this section is formally equivalent to the n-player version of the analysis of Section 4 where the designer chooses $\alpha_i \in \{0,1\}$ $\forall i$.

[17]In this region one could think of excluding the high or the medium type contestants in order to make the low type willing to participate. However, this is never optimal in either contests.

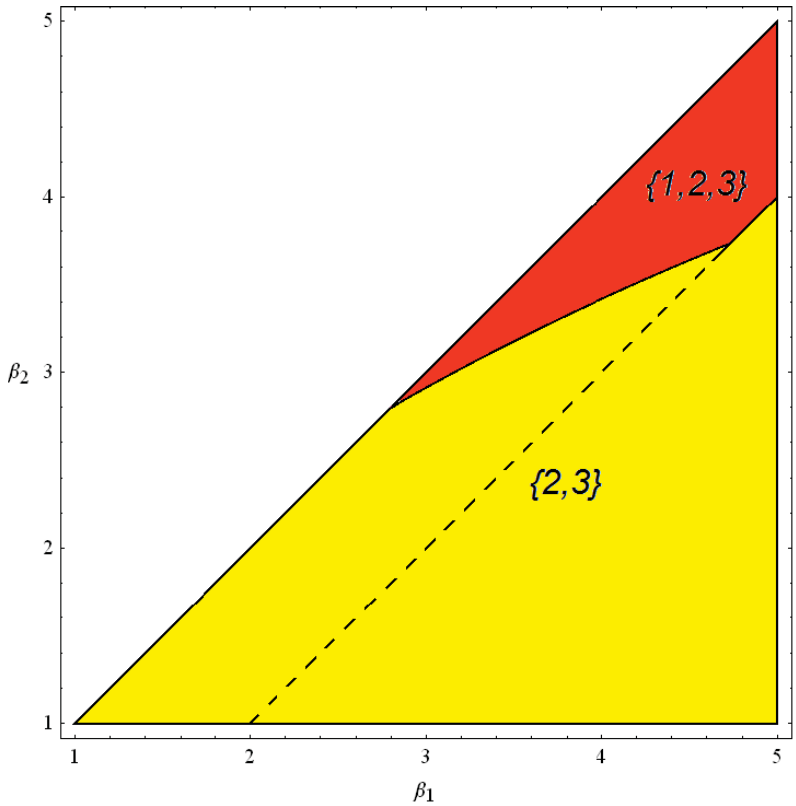

Figure 1: Optimal set of finalists in a 3-contestant quality contest.

abilities of winning become roughly $\frac{1}{2}$ instead of $\frac{1}{3}$, and thus each of them exerts more effort: that is, exclusion increases e_i. Since $\widehat{ewe} \approx e_i$, the designer benefits from excluding a contestant. For this reason we find that in this region it is optimal to exclude the low type, even though she would be willing to exert positive effort if selected, and thus we have a {2,3}-final. Note that the lack of effort by the excluded contestant in this region is not detrimental to $\widehat{ewe}$. On the contrary, in a stimulative contest, the lack of effort by the excluded contestant would be detrimental to $\widehat{sum}$, and in fact this negative effect of exclusion turns out to prevail on any other positive effect of exclusion on individual efforts in stimulative contests. This is why Fang (2002) finds that no exclusion is beneficial to stimulative contests - that is, the whole region between the two parallel lines would be completely red in a stimulative contest.

Proposition 4 *[Fang (2002)] In a stimulative contest with $\alpha_i = 1$ $\forall i$, and $\beta_1 \geq \beta_2 \geq ... \geq \beta_n > 0$ exclusion of contestants is not beneficial.*

From the 3-contestant example we see that for a wide variety of parameters, only medium and high types participate in the contest (yellow region). This happens either because the low type is too weak to be willing to participate (yellow region below the dotted line), or because the designer of the quality contest is better off excluding the low type from the contest (yellow region above the dotted line).

If we do the same exercise with four contestants, the conclusion is similar. Set $4 = \beta_1 \geq \beta_2 \geq \beta_3 \geq \beta_4 = 1$. The result is in Figure 2. Again, along the dotted

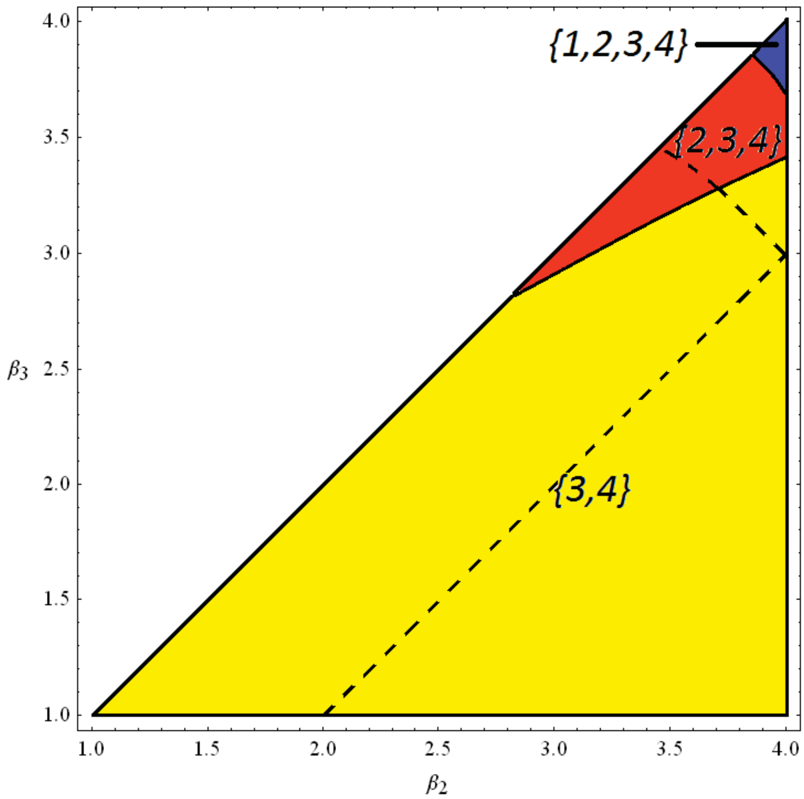

Figure 2: Optimal set of finalists in a 4-contestant quality contest.

lines a different degree of exclusion is needed to reach the final set of contestants. In particular, within a region of a certain color, the more we move upwards, the more exclusion is required to reach the final set of contestants. For instance, the finalists set {3,4} in the three yellow regions is achieved with different exclusion policies: in the big yellow triangle neither contestant 1 nor contestant 2 are willing to exert positive effort, in the yellow irregular pentagon contestant 1 is excluded and contestant 2 is not willing to exert effort, and in the small yellow triangle contestant 1 and contestant 2 are both excluded.

Extending the graphical analysis to more contestants would either require to add dimensions to the graph, or to give arbitrary values to all but two β_i's. In the former case we lose visualizability, in the latter we lose generality of the analysis. However, profitability of exclusion in a quality contest carries over. The intuition learnt in the 3-contestant and 4-contestant examples above is that exclusion is particularly profitable when contestants are of sufficiently homogenous types. This is the idea behind the proof of Proposition 5.

Proposition 5 *In a quality contest with $\alpha_i = 1$ $\forall i$, for any $n \geq 3$ there exists an open set containing $\{\beta_1, ..., \beta_n\}$ with $\beta_1 \geq \beta_2 \geq ... \geq \beta_n > 0$ such that exclusion of contestant(s) $\{1, ..., j\}$ with $1 \leq j \leq n-2$ is strictly beneficial (i.e., expected winning effort strictly increases).*

6 Comparative statics on contestants' types

A conclusion of the previous section is that for a large set of contestants' types the optimal number of finalists is 2. Then, it seems natural to run comparative statics on types in a two-contestant quality contest, which we do in this section. We use the tools developed so far to pin down the different channels through which changes of contestants' types affect first individual efforts, and second $\widehat{sum}$ and $\widehat{ewe}$. In order to gradually build intuition, it is crucial to distinguish between the playing field effect and the individual effect of changes of β_i.[18] In the two tables below we weaken contestant i, who is the underdog or the favorite.[19]

Weakening the underdog increases the asymmetry between contestants, and hence the playing field effect is negative for both contestants (the favorite is more favorite, and the underdog is less likely to win). Moreover, the effort of the underdog drops further, because her marginal cost of effort increases - i.e., negative individual effect.

$\uparrow \beta_i$ and i is the **underdog** $\rightarrow$

	playing field effect	individual effect	All in all
e_i	$\downarrow$	$\downarrow$	$e_i \downarrow$
e_j	$\downarrow$		$e_j \downarrow$

Weakening the favorite decreases the asymmetry between contestants, and hence the playing field effect is positive for both contestants (more levelled playing field). However, the effort of the favorite undergoes a negative individual effect, because her marginal cost of effort increases. All in all, the negative individual effect overcomes the positive playing field effect, and thus the effort of the favorite decreases as she is weakened.

$\uparrow \beta_i$ and i is the **favorite** $\rightarrow$

	playing field effect	individual effect	All in all
e_i	$\uparrow$	$\downarrow$	$e_i \downarrow$
e_j	$\uparrow$		$e_j \uparrow$

What are the final effects on $\widehat{sum}$ and $\widehat{ewe}$ of increasing β_i? The only straightforward conclusion from the above analysis is that $\widehat{sum}$ decreases when the underdog is weakened, because both contestants' efforts decrease. The other comparative statics depend on values of (α, β) and are contained in Proposition 6 and Proposition 7. We first report the results, and then discuss the intuition behind them.[20]

[18]The meaning of playing field and individual effects can be seen in (5). Note that changes in α_i's as in Section 4 have only playing field effects, whereas changes in β_i's have also individual effects. This is why we first analyzed changes in α_i's, and now in β_i's, and we keep these two analyses separate.

[19]Comparative statics on individual efforts are straightforward from (5), and thus proofs are omitted; we rather focus on results and intuitions. Comparative statics on $\widehat{sum}$ and $\widehat{ewe}$ are proved in Proposition 6 and Proposition 7.

[20]Despite the comparative statics on individual efforts are well-known (see Malueg and Yates (2005)), to the best of our knowledge we are the first to provide comparative statics on $\widehat{sum}$, and thus in case of weakening of the underdog to highlight the trade-off between positive playing field effect and negative individual effect.

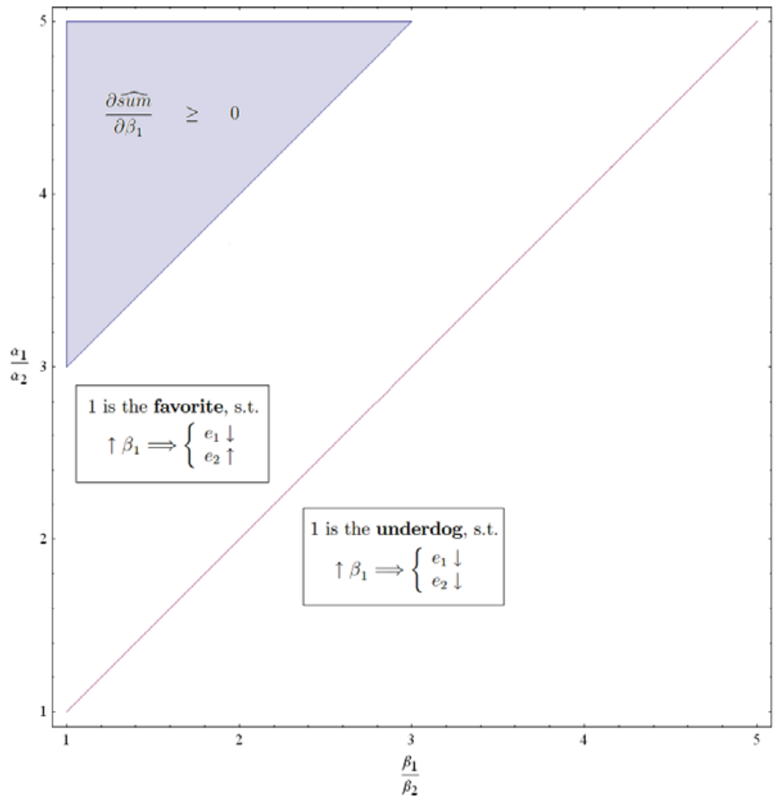

Figure 3: Region where weakening contestant 1 is beneficial to $\widehat{sum}$.

Proposition 6 *Weakening the underdog is not beneficial to $\widehat{sum}$.*
Weakening the favorite is beneficial to $\widehat{sum}$ iff

$$\beta \leq \alpha - 2 \tag{10}$$

where contestant 1 is the favorite.

Proposition 7 *Weakening contestant 1, who is either the favorite or the underdog, is beneficial to $\widehat{ewe}$ iff*

$$3\alpha + \beta^2 - 2\alpha\beta \leq 0 \tag{11}$$

Propositions 6 and 7 convey the main message of this section: while a weakening of the favorite could be beneficial both in a quality contest and in a stimulative contest, a weakening of the underdog could be beneficial only in a quality contest. We plot conditions (10) and (11) in the (α, β)-space in order to have a graphical support when building intuition. Above the 45° line contestant 1 - who is weakened in the propositions above - is the favorite, and below the 45° line contestant 1 is the underdog. We rewrite in two boxes in the figures the implications on efforts discussed above. Furthermore, assume $\beta \geq 1$, so that 1 (2) is the high (low) type.[21]

[21] Assuming $\beta \geq 1$ is wlog since we can swap i and j's identities, and retrieve the same qualitative graph in the $(\alpha, \beta) \in [0, \infty) \times [0, 1]$ space, rather than in the $(\alpha, \beta) \in [0, \infty) \times [1, \infty)$ space. For instance, condition (10), $\beta \leq \alpha - 2$, if we swap identities, reads $\frac{1}{\beta} \leq \frac{1}{\alpha} - 2$, which could be depicted in the $(\alpha, \beta) \in [0, \infty) \times [0, 1]$ space, but would correspond to the one we depict in the $(\alpha, \beta) \in [0, \infty) \times [1, \infty)$ space.

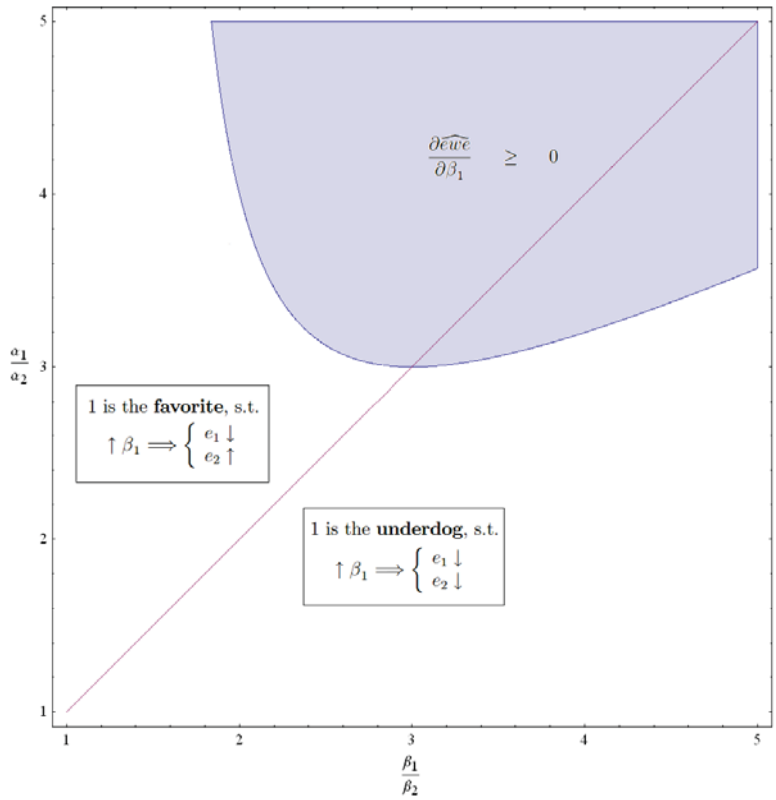

Figure 4: Region where weakening contestant 1 is beneficial to $\widehat{ewe}$.

This has a clear implication to keep in mind: $e_1^* \leq e_2^*$ (see (5)).

Figure 3. Weakening the underdog (below the 45° line) implies $e_1 \downarrow$ and $e_2 \downarrow$, and thus $\widehat{sum}$ decreases. On the other hand, weakening the favorite (above the 45° line) implies $e_1 \downarrow$ and $e_2 \uparrow$, thus $\widehat{sum}$ might increase or decrease. As explained above, this is due to the positive playing field effect and the negative individual effect. The former should prevail over the latter in order to have that $\frac{\partial \widehat{sum}}{\partial \beta_1} \geq 0$. The playing field effect is stronger the more unlevelled is the playing field: that is, levelling the playing field becomes the designer's "priority" in that it is the main driver of efforts.[22] Very unlevelled playing fields occur when sufficiently far from $S = 1$ (i.e., the 45° line), and this is why $\frac{\partial \widehat{sum}}{\partial \beta_1} \geq 0$ in the top-left area of Figure 3, corresponding to (10).

Figure 4. The region (11) is partially above and partially below the 45° line - that is, both weakening the underdog or the favorite could increase $\widehat{ewe}$. Weakening of the favorite yields $\overset{\downarrow}{p_1}\overset{\downarrow}{e_1} + \overset{\uparrow}{p_2}\overset{\uparrow}{e_2}$, and weakening of the underdog yields $\overset{\downarrow}{p_1}\overset{\downarrow}{e_1} + \overset{\uparrow}{p_2}\overset{\downarrow}{e_2}$, thus it is intuitive that the region where the former is beneficial (above the 45° line) is bigger than the region where the latter is beneficial (below the 45° line). Consider these two cases separately. First, weakening of the favorite ($\overset{\downarrow}{p_1}\overset{\downarrow}{e_1} + \overset{\uparrow}{p_2}\overset{\uparrow}{e_2}$) is beneficial unless p_1, which is the weight of the low and decreasing effort e_1, is sufficiently high to overcome the positive effect of $\overset{\uparrow}{e_2}$. Now, p_1 is sufficiently high in points sufficiently

[22] In fact, $\lim_{S \to 0} e_i^* = \lim_{S \to \infty} e_i^* = 0$ for $i = 1, 2$ regardless of β_1 and β_2. See (5).

close to the vertical axis.[23] This is why the region where $\frac{\partial \widehat{ewe}}{\partial \beta_1} \geq 0$ is qualitatively sufficiently far from the vertical axis. Second, weakening of the underdog $(\overset{\downarrow}{p_1}\overset{\downarrow}{e_1}+\overset{\uparrow}{p_2}\overset{\downarrow}{e_2})$ is rather thorny. Remember that $e_1 \leq e_2$, so if there is sufficient difference between e_2 and e_1, the positive effect of $\overset{\uparrow}{p_2}$, prevails on the negative $\overset{\downarrow}{e_1}$ and $\overset{\downarrow}{e_2}$.[24] Sufficient difference between e_2 and e_1 is achieved when β is sufficiently far from 1. In fact, a necessary condition for (11) is that $\beta \geq 3$,[25] which implies $e_2 \geq 3e_1$ - see (5). At the same time, now the playing field effect is negative (the underdog is further less likely to win), and thus the argument behind Figure 3, namely to avoid very unlevelled playing field, carries over. This is why this region is also sufficiently far from the horizontal axis.[26] In fact, a necessary condition for (11) is that $S \geq \frac{1}{2}$.[27] We provide a numerical example consistent with the above reasoning; if types are sufficiently heterogenous and the playing field sufficiently levelled, weakening the underdog is beneficial in a quality contest.

Example 8 *Consider* $(\beta_1, \beta_2) = (5, 1)$ *and* $S = 4/5$, *so that: i) types are sufficiently heterogenous, ii) the playing field sufficiently levelled, and iii) contestant 1 is the underdog.*[28] *Then,* $\widehat{ewe} = \frac{116}{729}$, *see (8).*

Now, consider $\beta_1 = 6$ *(weakening of the underdog), and keep the rest unchanged. Then,* $\widehat{ewe} = \frac{4}{25} > \frac{116}{729}$. *Thus, weakening the underdog yielded an increase in* $\widehat{ewe}$. *In fact, this numerical example lies in the region of Figure 4 below the* $45°$ *line where* $\frac{\partial \widehat{ewe}}{\partial \beta_1} \geq 0$.

7 Conclusions

In a standard model of lottery contest we propose that the contest designer benefits from the expected winning effort. We compare the optimal contest design in this

[23]Recall that $S = \alpha/\beta$, and thus $S = 0$ along the horizontal axis, $S = 1$ along the $45°$ line, and $S \to \infty$ along the vertical axis. Also, from (6) we know $p_1^* = \frac{S}{S+1}$.

[24]Start from $\widehat{ewe} = p_1 e_1 + p_2 e_2$, with $e_1 \leq e_2$. Increase p_2 by δ and decrease both efforts by a (this decrease needs not to be symmetric). Then for $\widehat{ewe}$ to increase, we need

$$\begin{aligned} p_1 e_1 + p_2 e_2 \quad &\leq \quad (p_1 - \delta)(e_1 - a) + (p_2 + \delta)(e_2 - a) \\ &\iff \quad 0 \leq -p_1 a - \delta(e_1 - a) - p_2 a + \delta(e_2 - a) \\ &\iff \quad 0 \leq -a - \delta(e_1 - e_2) \\ &\iff \quad \delta \geq \frac{-a}{e_1 - e_2} \end{aligned}$$

which holds if δ is sufficiently large as opposed to the efforts' drop relatively to efforts' difference. δ being sufficiently large means that p_2 increases sufficiently.

[25]$3\alpha + \beta^2 - 2\alpha\beta \leq 0 \iff 3\alpha - \alpha\beta + \beta^2 - \alpha\beta \leq 0 \iff \alpha(3 - \beta) + \beta(\beta - \alpha) \leq 0$. Since $\beta \geq \alpha$, it has to be that $\beta \geq 3$.

[26]Note that the argument that, if the playing field is too unlevelled, the playing field issue becomes the priority, is consistent with Corollary 3: never leave the playing field too unlevelled.

[27](11) is equivalent to $3\alpha + \beta(\beta - 2\alpha) \leq 0$. Since $\alpha \geq 0$, it has to be that $\beta \leq 2\alpha$, or equivalently that $S \geq 1/2$.

[28]In fact, when $(\beta_1, \beta_2) = (5, 1)$, any $\alpha \leq 5$ guarantees that contestant 1 is the underdog. We set $\alpha = 4$, so that $S = 4/5$, consistently with the need for a sufficiently levelled playing field in order for the result to hold.

scenario (quality contest) as opposed to the conventional case where the designer benefits from every contestant's effort equally (stimulative contest). We find that, on the contrary of a stimulative contest designer, a quality contest designer may benefit from: unlevelled playing field, exclusion of weak contestants, and weakening of the underdog. The results of this paper are a first step meant to stress the importance of specifying the objective function of the contest designer, as well as to pave the way for future works in this direction.

A Appendix A: Proofs

Proof of Proposition 2. The first-order condition of $\max_{\alpha} \widehat{ewe}$ is

$$\begin{aligned}
\frac{\partial \widehat{ewe}}{\partial \alpha} &= 0 \\
\iff & \left[\frac{1-2S}{(1+S)^4} * \left(\frac{S}{\beta_1} + \frac{1}{\beta_2}\right) + \frac{S}{(1+S)^3 \beta_1}\right] \frac{1}{\beta} = 0 && (12) \\
\iff & \frac{1-2S}{1+S} \left(\frac{S}{\beta_1} + \frac{1}{\beta_2}\right) + \frac{S}{\beta_1} = 0 \\
\iff & (1-2S)(S+\beta) + S(S+1) = 0 \\
\iff & S^2 + 2(\beta - 1)S - \beta = 0 && (13)
\end{aligned}$$

One root of (13) is negative, and the positive one is

$$S^q = 1 - \beta + \sqrt{(1-\beta)^2 + \beta}$$

which corresponds to (9). S^q is the unique maximum of (8) because $\lim_{S \to 0} \widehat{ewe} = \lim_{S \to \infty} \widehat{ewe} = 0$, $\widehat{ewe}$ is continuously differentiable in S, (12) is strictly positive when $S = 0$, and S^q is the unique positive root. ■

Proof of Corollary 3. We first prove that $\alpha^q \in [\frac{1}{2}\beta, 2\beta]$.

$$\begin{aligned}
& \alpha^q \leq 2\beta^r \\
& \iff 1 - \beta + \sqrt{(1-\beta)^2 + \beta} \leq 2 \\
& \iff (1-\beta)^2 + \beta \leq (1+\beta)^2 \\
& \iff \beta \leq 4\beta
\end{aligned}$$

$$\begin{aligned}
& \alpha^q \geq \frac{1}{2}\beta^r \\
& \iff 1 - \beta + \sqrt{(1-\beta)^2 + \beta} \geq \frac{1}{2} \\
& \iff 2\sqrt{(1-\beta)^2 + \beta} \geq 2\beta - 1 && (14)
\end{aligned}$$

If $\beta < \frac{1}{2}$, (14) holds trivially. If $\beta \geq \frac{1}{2}$, (14) is equivalent to $4\left[(1-\beta)^2 + \beta\right] \geq [2\beta - 1]^2 \iff 4\left[\beta^2 - \beta + 1\right] \geq 4\beta^2 - 4\beta + 1$ which also holds. The boundaries of S trivially follows, and the boundaries for p_i's follow from (6). ■

Proof of Proposition 5. First, it is well known (see for instance Fullerton and McAfee (1999)) that equilibrium effort is

$$e_i^* = \frac{m-1}{\sum_{j \in m} \beta_j} \left[1 - \frac{\beta_i (m-1)}{\sum_{j \in m} \beta_j}\right] \tag{15}$$

when m contestants exert positive effort. We denote the set of m contestants by $M=\{n-m,...,n\}$.

From (15),

$$\begin{aligned}
p_i^* &= \frac{\left[1-\frac{\beta_i(m-1)}{\sum\limits_{j\in M}\beta_j}\right]}{\sum\limits_{k\in M}\left[1-\frac{\beta_k(m-1)}{\sum\limits_{j\in M}\beta_j}\right]} \\
&= \frac{\sum\limits_{j\in m}\beta_j-\beta_i(m-1)}{m\sum\limits_{j\in M}\beta_j-(m-1)\sum\limits_{j\in M}\beta_j} \\
&= \frac{\sum\limits_{j\in M}\beta_j-\beta_i(m-1)}{\sum\limits_{j\in M}\beta_j} \\
&= 1-\frac{\beta_i(m-1)}{\sum\limits_{j\in M}\beta_j}
\end{aligned}$$

We can now provide a general formula for the expected winning effort given that m contestants exert positive effort, which we denote $\widehat{ewe}_m$

$$\begin{aligned}
\widehat{ewe}_m &= \sum_{i\in M}p_i^*e_i^* \\
&= \frac{m-1}{\sum\limits_{j\in M}\beta_j}\sum_{i\in M}\left[1-\frac{\beta_i(m-1)}{\sum\limits_{j\in M}\beta_j}\right]^2 \\
&= \frac{m-1}{\sum\limits_{j\in M}\beta_j}\left[m+(m-1)^2\frac{\sum\limits_{i\in M}\beta_i^2}{\left(\sum\limits_{j\in M}\beta_j\right)^2}-2(m-1)\frac{\sum\limits_{i\in M}\beta_i}{\sum\limits_{j\in M}\beta_j}\right] \\
&= \frac{m-1}{\sum\limits_{j\in M}\beta_j}\left[m+(m-1)^2\frac{\sum\limits_{i\in M}\beta_i^2}{\left(\sum\limits_{j\in M}\beta_j\right)^2}-2(m-1)\right] \\
&= \frac{m-1}{\sum\limits_{j\in M}\beta_j}\left[2-m+(m-1)^2\frac{\sum\limits_{i\in M}\beta_i^2}{\left(\sum\limits_{j\in M}\beta_j\right)^2}\right] \qquad (16)
\end{aligned}$$

Second, the intuition behind the 3-contestant and 4-contestant examples suggests to look for profitable exclusion under homogeneity of types. Thus, start from $\beta_1=$

$\beta_2 = ... = \beta_n = \bar{\beta}$. It is trivial to see that $e_i^* \geq 0$ for all contestants - see (15). We now prove that $\widehat{ewe}_{m-1} > \widehat{ewe}_m$, applying formula (16),

$$\begin{aligned}
\widehat{ewe}_{m-1} \quad &> \quad \widehat{ewe}_m \\
&\iff \frac{m-2}{(m-1)\bar{\beta}}\left[3-m+(m-2)^2\frac{(m-1)\bar{\beta}^2}{(m-1)^2\bar{\beta}^2}\right] > \frac{m-1}{m\bar{\beta}}\left[2-m+(m-1)^2\frac{m\bar{\beta}^2}{m^2\bar{\beta}^2}\right] \\
&\iff \frac{m-2}{(m-1)^2}\left[(3-m)(m-1)+(m-2)^2\right] > \frac{m-1}{m^2}\left[(2-m)m+(m-1)^2\right] \\
&\iff \frac{m-2}{(m-1)^2} > \frac{m-1}{m^2} \\
&\iff m^2(m-2) > (m-1)^3 \\
&\iff (m-1)^2 > m
\end{aligned}$$

which holds true for all $m \geq 3$. Since $\widehat{ewe}$ is continuous in β_i's, for sufficiently small changes in the vector of β's the inequality $\widehat{ewe}_{m-1} > \widehat{ewe}_m$ continues to hold. This complete the proof. ■

Proof of Proposition 6. Consider $\frac{\partial \widehat{sum}}{\partial \beta_1}$, where contestant 1 is either the favorite or the underdog. In the following steps we use $\frac{\partial S}{\partial \beta_1} = -\frac{S}{\beta_1}$ and $S = \frac{\alpha}{\beta}$.

$$\begin{aligned}
\frac{\partial \widehat{sum}}{\partial \beta_1} \quad &\geq \quad 0 \\
&\iff \frac{\partial}{\partial \beta_1}\left[\frac{S}{(1+S)^2}\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right)\right] \geq 0 \\
&\iff \frac{1-S}{(1+S)^3}\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right)\left(-\frac{S}{\beta_1}\right) - \frac{S}{(1+S)^2}\frac{1}{\beta_1^2} \geq 0 \\
&\iff \frac{1-S}{1+S}(1+\beta)+1 \leq 0 \qquad (17) \\
&\iff \frac{\beta-\alpha}{\beta+\alpha}(1+\beta)+1 \leq 0 \\
&\iff (\beta-\alpha)(1+\beta) \leq -\alpha-\beta \\
&\iff 2\beta+\beta^2-\alpha\beta \leq 0 \\
&\iff \beta \leq \alpha - 2 \qquad (18)
\end{aligned}$$

Suppose $S \leq 1$, such that contestant 1 is the underdog. The right-hand side of (17) is positive, and thus $\frac{\partial \widehat{sum}}{\partial \beta_1} \leq 0$. That is, weakening the underdog is detrimental to $\widehat{sum}$. Suppose $S \geq 1$, such that contestant 1 is the favorite. Then $\alpha \geq \beta$, and condition (18) might or might not hold. ■

Proof of Proposition 7. In the following we use $\frac{\partial S}{\partial \beta_1} = -\frac{S}{\beta_1}$ and $S = \frac{\alpha}{\beta}$.

$$\begin{aligned}
\frac{\partial \widehat{ewe}}{\partial \beta_1} &\geq 0 \\
&\iff \frac{\partial}{\partial \beta_1}\left[\frac{S}{(1+S)^3}\left(\frac{S}{\beta_1}+\frac{1}{\beta_2}\right)\right] \geq 0 \\
&\iff \frac{1-2S}{(1+S)^4}\left(\frac{S}{\beta_1}+\frac{1}{\beta_2}\right)\left(-\frac{S}{\beta_1}\right)+\frac{S}{(1+S)^3}\left(\frac{(-S/\beta_1)\beta_1 - S}{\beta_1^2}\right) \geq 0 \\
&\iff \frac{1-2S}{1+S}(S+\beta)+2S \leq 0 \\
&\iff \frac{\beta-2\alpha}{\beta+\alpha}\left(\alpha+\beta^2\right)+2\alpha \leq 0 \\
&\iff \alpha\beta+\beta^3-2\alpha\beta^2 \leq -2\alpha\beta \\
&\iff 3\alpha+\beta^2-2\alpha\beta \leq 0 \qquad (19)
\end{aligned}$$

Note that condition (19) could hold or not regardless of whether $S \gtreqless 1$. That is, weakening of underdog or favorite could be beneficial to $\widehat{ewe}$. ■

B Appendix B: noisy contest vs non-noisy contest

We analyzed a model where contestants compete in a *noisy* contest. However, one might wonder to what extent the designer should be concerned with the presence of the noise. In this Appendix we analyze $\widehat{sum}$ and $\widehat{ewe}$ in contests with noise (Tullock contest, or CSF), and in contests without noise (all-pay auction, or APA). All α_i's equal 1, and $n = 2$. In APA, the equilibrium is in mixed-strategies. This makes the analytical derivation of the result rather space-demanding. Thus, we only deliver the intuition in what follows, and a formal proof is available upon request.

Figure 5. If $\beta_1 = \beta_2$, competition between the two contestants is maximum, and thus the noise can only be detrimental. In other words, the greater is the noise, the less their effort is important in determining who wins, the less effort they exert. Hence, sufficiently close to $\beta_1 = \beta_2$ APA is optimal. If the two contestants have very heterogenous types (β_2/β_1 is close to 0), then the noise is desirable because it gives hope to the low type to win, and this is beneficial to competition. Hence, sufficiently close to $\beta_2/\beta_1 = 0$ CSF is optimal. These two extreme cases are common to stimulative and quality contests. Now, the threshold where the designer is indifferent between APA and CSF is lower for a quality contest than for a stimulative contest - that is, a quality contest designer benefits more from APAs than a stimulative contest designer. The reason is as follows. In APA contestants mix over the same support, and the unique positive mass point in their equilibrium mixed strategy is in the strategy of the low type to exert 0-effort. This equilibrium probability of exerting 0-effort for the low type increases in the asymmetry of types. A mass in 0 is detrimental to stimulative contests, since the designer equally benefits from efforts of the high type and of the low type, but it is desirable in a quality contest;

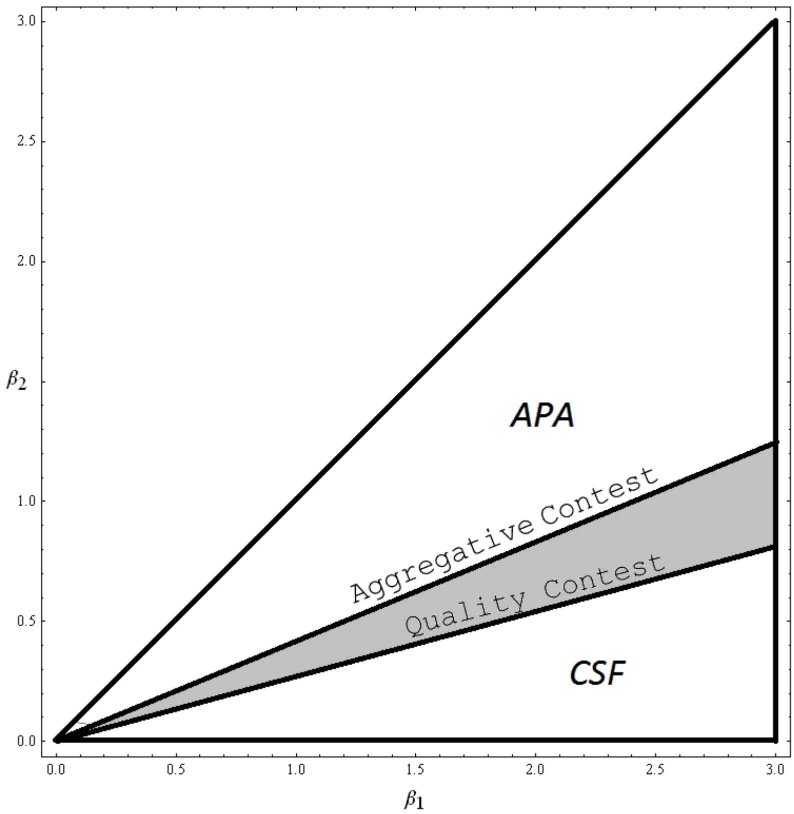

Figure 5: APA vs lottery CSF in stimulative and quality contest.

while a vector of efforts $\{10, 0\}$ yields $\widehat{ewe} = 10$, and a vector of efforts $\{10, x\}$ with $x \in (0, 10)$ yields $\widehat{ewe} < 10$. In words, a small but positive effort severely drags down the expected winning effort, on the contrary of a 0 effort; a quality contest designer does not want to risk ending up with the low type's project winning the contest.

References

[1] Andreff, W., and Szymanski, S., 2006. "Handbook On The Economics Of Sport", Edward Elgar Publishing Limited.

[2] Azmat, G., Möller, M., 2009. "Competition Amongst Contests". *The RAND Journal of Economics*, 40, 743-768

[3] Baye, M. R., Kovenock, D., and de Vries, C. G., 1993. "Rigging the Lobbying Process: An Application of the All-Pay Auction". *The American Economic Review*, 83, 289-294

[4] Brown, J., 2001. "Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars". *Journal of Political Economy*, 119, 982-1013

[5] Cohen, C., Kaplan, T. R., and Sela, A, 2008. "Optimal Rewards in Contests". *The RAND Journal of Economics*, 39, 434-451

[6] Drugov, M., and Ryvkin, D., 2016. "Biased Contests for Symmetric Players". *Games and Economic Behavior*, forthcoming

[7] Falconieri, S., Palomino, F., and Sákovics, J., 2004. "Collective versus individual sale of television rights in league sports". *Journal of the European Economic Association,* 2, 833–862

[8] Fang, H., 2008. "Lottery versus all-pay auction models of lobbying". *Public Choice*, 112, 351–371

[9] Franke, J., 2012. "Affirmative action in contest games". *European Journal of Political Economy*, 28, 105-118

[10] Franke, J., Kanzow, C., Leininger, W., and Schwartz, A. 2013. "Effort maximization in asymmetric contest games with heterogeneous contestants". *Economic Theory*, 52, 589-630

[11] Jonsson, S., and Schmutzler, A., 2015. "All-pay Auctions: Implementation and Optimality". Working Paper

[12] Jia, H. 2008. "A stochastic derivation of the ratio form of contest success functions". *Public Choice*, 135, 125-130

[13] Malueg, D. A., and Yates, A. J., 2005. "Equilibria and comparative statics in two-player contests". *European Journal of Political Economy,* 21, 738–752

[14] Moldovanu, B., and Sela, A., 2001. "The Optimal Allocation of Prizes in Contests". *The American Economic Review*, 91, 542-558

[15] Moldovanu, B., Sela, A., and Shi, X., 2007 "Contests for Status". *Journal of Political Economy*, 115, 338-363

[16] Nti, K.O., 2004. "Maximum efforts in contests with asymmetric valuations". *European Journal of Political Economy,* 20, 1059–1066

[17] Palomino, F., and Sákovics, J., 2004. "Inter-league competition for talent vs. competitive balance". *International Journal of Industrial Organization,* 22, 783-797

[18] Runkel, M., 2006. "Total effort, competitive balance and the optimal contest success function". *European Journal of Political Economy*, 22, 1009-1013

[19] Segev, E., and Sela, A., 2014. "Multi-stage sequential all-pay auctions". *European Economic Review,* 70, 371–382

[20] Vrooman, J., 2012. "Two to Tango: Optimum Competitive Balance in Pro Sports Leagues". In The Econometrics of Sport, Palacido Rodriguez, Stefan Kesenne, and Juame Garcia eds. Edward Elgar.

Chapter 3

Costly Voting under Complete Information

1 Introduction

Pivotal voter models with costly voting started under complete information with the seminal contribution of Palfrey and Rosenthal (1983) (henceforth, PR), where two groups of individuals each preferring one of two alternatives simultaneously decide between abstaining or voting for their preferred alternative. The winner is decided by majority rule. Technical difficulties and multiplicity issues allowed them to analyze only special cases.[1] Two years later, the same authors proposed in Palfrey and Rosenthal (1985) to drop the assumption of complete information, and from then onwards the literature has almost exclusively focused on private information on the cost of voting. Some of the most relevant contributions which analyze private-information costly-voting pivotal voter models are Campbell (1999), Borgers (2004), Feddersen and Sandroni (2006), and Taylor and Yildirim (2015).[2] We believe the complete information setting also deserves attention, hence the goal of this paper is to complement the existing private information literature, and to move the PR analysis a step forward.

The literature has shown that assuming private information on the cost of voting typically has technical advantages. In particular the equilibrium is unique and the strategies are completely characterized by a single cut-off value: supporters of A (B) vote if and only if the cost of voting is lower than a threshold c_A (c_B).[3] As PR showed, uniqueness and cut-off strategies are not generally present in models of complete information. Nevertheless, we first show that, if the cost of voting is sufficiently high at least for the supporters of one of the two alternatives, the equilibrium is unique. We fully characterize it. If instead the cost of voting is sufficiently low for all individuals, we characterize three classes of equilibria, and show that any equilibrium must belong to one of these three classes, regardless of the number of individuals. Furthermore, we propose a novel equilibrium refinement that always singles out a unique equilibrium. This refinement says that the equilibrium probability of voting is continuous in the cost of voting. In fact, it would be hard to claim that negligible changes in the cost of voting could bring about drastic changes in the probabilities of voting. For example, if the voting center station moves slightly away from the home of an individual, her probability of voting also changes negligibly. The unique equilibrium pinned down by the continuity refinement is proved to belong always to the same class of equilibria out of the three classes previously characterized. The features of the equilibrium that we analyze are as follows. First of all, we find a

[1]In particular, PR's setting is one of identical individual benefits from winning across groups, and they analyze: i) identical group sizes and symmetry of strategies across groups, and ii) aggregate probabilities of voting across individuals of different groups summing to 1. Besides these special cases of our analysis, they also analyze two different settings: one where there is a status quo (ties are broken in favor of one group, instead of randomly), and one, namely "k equilibria", where individuals of one group mix with identical probability, whereas among the individuals of the other group, k vote with probability 1 and the remaining with probability 0.

[2]Another way to exploit private information to simplify the analysis is provided in Myerson (1998,1998,2000). He suggests an alternative model in which the size of the electorate is a Poisson random variable. As Krishna and Morgan (2012, JET) claim, Myerson's approach "has the important advantage of considerably simplifying the analysis of pivotal events". In Myerson's games, citizens' preferences are determined via a *privately* observed stochastic draw.

[3]Note that all papers cited in the main text of the previous paragraph assume private information and have only equilibria in cut-off strategies, but PR (83).

turnout upper-bound: the sum of the equilibrium probabilities of voting of an m-individual and an n-individual is less than 1. Moreover, members of the majority group with higher cost-to-benefit ratio have higher probability of being pivotal in equilibrium, and thus –intuitively– vote with a higher probability in equilibrium. This carries over even if the two groups are asymmetric only in the cost-to-benefit ratio and symmetric in size. If instead the two groups are symmetric in cost-to-benefit ratio and asymmetric in size, members of the minority group vote with a strictly higher probability than those in the majority do. This latter result already exists in models of private information on the cost of voting, see Taylor and Yildirim (2010). The fact that members of the minority vote more than members of the majority is called the "underdog effect".[4] However, we find that if the minority has a sufficiently higher cost-to-benefit ratio than the majority, the underdog effect disappears (i.e., the members of the minority group votes with lower probability than members of the majority group). This result is trivial per se (infinite cost-to-benefit ratio yields necessary 0 probability of voting) unless something is said about the asymmetry in cost-to-benefit needed to break the underdog effect: we show that if the ratio of cost-to-benefit ratios is greater than the number of members of the minority then the underdog effect is contradicted and the members of the minority group vote with 0 probability regardless of their (strictly positive) cost-to-benefit ratio. We furthermore perform comparative statics in the number of individuals, and give the intuition behind it, along with the construction of the equilibrium itself.

Our complete information setting springs from PR, however we depart from them in that an individual's benefit of having the favorite alternative win can be asymmetric according to whether the individual supports one or the other alternative, meaning that the personal benefit can differ between individuals supporting one alternative or the other. Our generalization of PR to asymmetric benefits is natural especially when the group sizes are asymmetric. For instance, think of an economics department consisting of several microeconomists and a few macroeconomists, all called to vote over who to hire between two job market candidates: a microeconomist and a macroeconomist.[5] Both types of economists are better-off if the newly hired candidate is of their same type. Furthermore, the benefit for a macroeconomist from hiring another macroeconomist is greater than the one for a microeconomist from hiring a microeconomist because of the asymmetric size of the two groups; that is, since macroeconomists are fewer, having another macroeconomist in the department sharply increases each macroeconomist's coauthoring possibilities, whereas the benefit for a microeconomist from having a new microeconomist in the department is lower because they are already plenty. In other words, the benefit is asymmetric across individuals of different groups. This asymmetry gives rise to asymmetric willingness to vote.[6]

[4] Laboratory experiments confirm the underdog effect, see Levine, D., Palfrey, T., 2007. The paradox of voter participation? A laboratory study. Amer. Polit. Sci. Rev. 101 (1), 1431 58.

[5] The cost of voting in this case is the opportunity cost of showing up to vote that day instead of, for example, being on vacation.

[6] Another example of asymmetric benefit across groups is the following. Residents of two neighborhoods are called to vote over the location of a new school in one of the two neighborhoods. In neighborhood 1 there is already a school, in neighborhood 2 there is none: thus, despite the fact that each resident strictly prefers the school to be located in her neighborhood, residents in neighborhood 1 care less than residents in neighborhood 2 about the location of the school since

2 Model

Consider a complete information setting where there are two groups of individuals of size m and n, with $m, n \in \mathbb{N}^+$. Throughout the paper we assume $m > n > 1$: the analysis of $n = 1$ is ruled out to avoid dealing with trivial cases, and the analysis of $m = n$ produces peculiar results and is left to Appendix C. We use sub-index $i \in \{m, n\}$ to identify the group with a slight abuse of notation. The individuals are called to cast a vote between two alternatives, M and N. An individual of group m prefers alternative M, and an individual of group n prefers alternative N. That is, if M wins, the payoff of an individual m increases by $\Delta\pi_m \in \mathbb{R}_{++}$, and similarly by $\Delta\pi_n \in \mathbb{R}_{++}$ for individuals n if N wins. Individuals choose whether to vote for their preferred alternative or to abstain, since voting for the non-preferred alternative is strictly dominated. If an individual casts a vote, she faces a group-specific cost of voting, $c_i \in \mathbb{R}_{++}$. Thus, the increase in payoff –net of cost of voting– for an individual i when her preferred policy wins is $\Delta\pi_i - c_i$ if she voted, and $\Delta\pi_i$ if she did not vote. Individuals vote simultaneously and the winning alternative is decided by majority rule. Ties are broken by a fair coin toss.

Each individual i chooses her probability of voting, denoted by p_i, that maximizes her expected payoff, given the choices of all other individuals. We consider Quasi-Symmetric Nash Equilibria (QSNE), that is, individuals of group i follow the same equilibrium strategy p_i^*. Besides being used in PR, the assumption of QSNE has been used in private-information pivotal-voter models to obtain that individuals adopt cut-off strategies. See for instance Börgers (2004) and Taylor and Yildirim (2010).

In a QSNE a pair (p_i^*, p_j^*) is an equilibrium if an individual of group $i \in \{m, n\}$ would not want to deviate from p_i^* if she expects every other individual of group i to also play p_i^* and all individuals of group j with $j \neq i$ to play p_j^*. A QSNE can be of one of the following three types:

1. "Pure-Pure": $(p_m^*, p_n^*) \in \{0, 1\}^2$
2. "Pure-Mix": $p_m^* \in \{0, 1\}, p_n^* \in (0, 1)$ or $p_m^* \in (0, 1), p_n^* \in \{0, 1\}$
3. "Mix-Mix": $(p_m^*, p_n^*) \in (0, 1)^2$.

Define A_i to be the probability that the vote of an individual of group i is pivotal. An individual of group i will cast a vote if:

$$A_i \Delta\pi_i \geq c_i$$

or

$$A_i \geq \frac{c_i}{\Delta\pi_i} \equiv B_i \tag{1}$$

for $i \in \{m, n\}$. The probabilities of being pivotal are defined as follows:[7]

$$A_i = \sum_{s=0}^{n} \binom{i-1}{s} \binom{j}{s} p_i^s (1-p_i)^{i-s-1} p_j^s (1-p_j)^{j-s} + \sum_{s=0}^{n-1} \binom{i-1}{s} \binom{j}{s+1} p_i^s (1-p_i)^{i-s-1} p_j^{s+1} (1-p_j)^{j-s-1} \tag{2}$$

there is already a school in neighborhood 1.

[7]Note that they already take into account that within groups individuals vote with the same probability.

for $i, j \in \{m, n\}$, $i \neq j$.

We now explain how the expression (2) is constructed. A single individual m, who computes her probability of being pivotal, takes as given the probabilities of voting (p_m, p_n) of all other individuals. The individual m is pivotal when her vote either breaks a tie, or when it creates one. In (2) the first summation is her probability of breaking a tie, and the second of creating a tie. She can break a tie with her vote, if the number of individuals that vote for m equals the number of individuals that vote for n. Call this number s. Out of $m-1$ other m-individuals, exactly s vote with probability $\binom{m-1}{s} p_m^s (1-p_m)^{m-s-1}$. On the other hand, out of n n-individuals, exactly s vote with probability $\binom{n}{s} p_n^s (1-p_n)^{n-s}$. The second summation of (2) is similarly constructed: individual m can create a tie with her vote, if the number of individuals that vote for m, (which is again called s), is one less than the number of individuals that vote for n.

3 Computing the equilibria

In this section we compute the QSNE of the voting game, given the relative costs of voting B_i and B_j, and classify them according to their type: "Pure-Pure", "Pure-Mix" or "Mix-Mix". We start with the first two types, ie. equilibria in which some individuals have pure strategies.

3.1 "Pure-Pure" and "Pure-Mix" equilibria

If $A_i < B_i$ or $A_i > B_i$, then individual i's dominant strategy is to abstain or to vote, respectively (i.e. pure strategy). Whereas when $A_i = B_i$, i is indifferent between voting and not (i.e. mixed strategy). B_i is therefore the minimum probability of being pivotal such that an individual i will vote. For that reason, an individual i whose B_i is greater than 1 does not vote in equilibrium.[8] We formalize this result in the following lemma.

Lemma 1. *If $B_i \geq 1$ then $p_i^* = 0$.*

Proof. Let $p_i^* > 0$. First, if $B_i > 1$, by (1) we have $A_i > 1$, which is a contradiction, since A_i is a probability. Second, if $B_i = 1$, by (1) we have $A_i = 1$. Then, we cannot have $p_i^* \in (0, 1)$ because of the following: if all individuals are randomizing between voting and not, then they cannot be pivotal with certainty, so $A_i \neq 1$ leading to a contradiction. Therefore we only need to rule out $B_i = p_i^* = 1$. For this we need to distinguish the following cases.

Case 1. If $p_m^* = 1$ all m individuals vote, which means that they win regardless of p_n^* because they are the largest group. But then, no n individual would want to face the cost of voting, $p_n^* = 0$. Thus, for a single m-individual deviation to $p_m = 0$ is profitable, leading to a contradiction.

Case 2. If $p_n^* = 1$ then in order to sustain $A_n = 1$ the n individuals must be certain that either n or $n-1$ of the m individuals vote. However, this can happen

[8]If the cost of voting for i is greater than the benefit of winning with certainty of being pivotal, in the unique equilibrium no member of i will vote (that is, $c_i > \Delta\pi_i$ contradicts $A_i \in [0, 1]$).

neither if $p_m^* \in (0,1)$ nor if $p_m^* = \{0,1\}$ (because this would imply m or 0 votes cast by group m). Thus, we reach a contradiction. $\square$

The previous lemma shows that if relative costs of voting are high enough for individuals of both groups, the only equilibrium that exists is the "Pure-Pure" one in which nobody votes.

Obviously, the situation described in Lemma 1 is not very interesting. Therefore, next we allow one of the two relative costs of voting to be low enough such that individuals from one of the two groups might consider voting, in other words $B_i \geq 1$ only for individuals i. Then Lemma 2 yields a simple and unique characterization of j's equilibrium strategy:

Lemma 2. *("Pure-Mix") For $B_i \geq 1$ and $B_j \in (0,1)$ the unique QSNE is that $p_i^* = 0$ and $p_j^* = 1 - B_j^{\frac{1}{j-1}}$, for all $i, j \in \{m, n\}$, $i \neq j$.*[9]

Proof. By Lemma 1, $p_i^* = 0$. Suppose $p_j^* = 0$. Then any single j individual would have an incentive to deviate and vote for sure in order to single-handedly decide the election in favor of the j-group. Thus $p_j = 0$ is not an equilibrium. On the other hand, suppose $p_j^* = 1$. This means that j group wins for sure with a margin of j votes. Then any single j-individual would have an incentive not to pay the cost without affecting the outcome. Thus $p_j = 1$ is not an equilibrium. Therefore $p_j^* \in (0,1)$. Plugging $p_i^* = 0$ in A_j we have $A_j = (1-p_j)^{j-1}$, and since (1) must hold with equality for individuals j to mix, we have:

$$(1 - p_j)^{j-1} = B_j$$

or equivalently:

$$p_j = 1 - B_j^{\frac{1}{j-1}}.$$

$\square$

For $x \in (0,1)$ the expression $1 - x^{\frac{1}{j-1}}$ is strictly decreasing in x. Therefore $p_j^* = 1 - B_j^{\frac{1}{j-1}}$ is strictly decreasing in c_j and strictly increasing in $\Delta\pi_j$. Higher individual cost-payoff ratio results in j-individuals voting with lower probability.

The two previous Lemmas examine cases in which individuals from at least one group find it too costly to vote, not matter what individuals from the other group are doing. These two cases gave rise to two types of equilibria "Pure-Pure" in which everybody's strictly dominant strategy is to not vote (Lemma 1), and "Pure-Mix" in which individuals from one group have a strictly dominant strategy to not vote and individuals from the other group mix (Lemma 2).

We examine next what happens when individuals from neither group have a strictly dominant strategy to abstain, ie. what happens when $B_m < 1$ and $B_n < 1$. Under these conditions individuals of both groups may vote with positive probability. This causes strategic interactions that may generate multiple equilibria.

It is easy to see that when $B_m < 1$ and $B_n < 1$ no "Pure-Pure" equilibria exist.

[9] As a reminder, in the beginning of this section we have assumed $m > n > 1$. It is easy to see that for $j = 1$ the unique equilibrium is: $p_i^* = 0$ and $p_j^* = 1$

Lemma 3. *For $B_i < 1$ and $B_j < 1$, no "Pure-Pure" QSNE exist, for all $i, j \in \{m, n\}$ and $i \neq j$.*

Proof. Assume $p_i^* = 0$. Then $p_j^* = 0$ cannot be a QSNE because $A_j = 1$ so deviation to voting for a j individual would be profitable. Also $p_j = 1$ cannot be a QSNE because a j-individual deviating to $p_j = 0$ would not affect the outcome of the election and save her cost of voting.

The only case left to analyze is $p_m^* = p_n^* = 1$. The n individuals lose for sure, and thus they would be better-off not to vote. □

After proving that for $B_m < 1$ and $B_n < 1$ no "Pure-Pure" equilibria exist, the next lemma establishes that for $B_m < 1$ and $B_n < 1$ a "Pure-Mix" equilibrium does exist. First, we need to define $\underline{B_i} = jB_j - (j-1)B_j^{\frac{j}{j-1}}$

Lemma 4. *For $B_i < 1$ and $B_j < 1$, there exists a "Pure-Mix" QSNE with $p_i^* = 0$ and $p_j^* = 1 - B_j^{\frac{1}{j-1}}$, for all $i, j \in \{m, n\}$ and $i \neq j$ if and only if $B_i \geq \underline{B_i}$.*

Proof. An equilibrium where $p_i^* = 0$ implies: $A_j = (1-p_j)^{j-1}$ and $A_i = (1-p_j)^j + jp_j(1-p_j)^{j-1}$. The former means that an individual j is pivotal only if none of her groupmates vote (her vote breaks the tie in which nobody votes). The latter means that an individual i is pivotal if none of j individuals vote or if only one of them votes. In order for the i-individuals to not want to vote we must have:

$$A_i \leq B_i,$$

or equivalently

$$(1-p_j)^j + jp_j(1-p_j)^{j-1} \leq B_i,$$

and similarly, for the j-group individual to mix we must have:

$$(1-p_j)^{j-1} = B_j, \tag{3}$$

dividing the two conditions and rearranging we get:

$$\begin{aligned} 1 - p_j + jp_j &\leq \frac{B_i}{B_j} \\ (j-1)p_j &\leq \frac{B_i}{B_j} - 1 \end{aligned} \tag{4}$$

Isolate p_j in (3) and plug it in (4) to get

$$(j-1)\left(1 - B_j^{\frac{1}{j-1}}\right) \leq \frac{B_i}{B_j} - 1 \tag{5}$$

Or equivalently,

$$-B_j^{\frac{j}{j-1}} \leq \frac{B_i - jB_j}{j-1} \tag{6}$$

$$B_i \geq jB_j - (j-1)B_j^{\frac{j}{j-1}} \equiv \underline{B_i}$$

□

$\underline{B_i}$ is an increasing bijection from $[0,1]$ to $[0,1]$, such that if $B_j = 0$, $\underline{B_i} = 0$, and if $B_j = 1$, $\underline{B_i} = 1$.

Note that the equilibria pinned down by Lemma 2 and Lemma 4 are essentially the same, the difference being that Lemma 2 provides the range of B_i's for which the equilibrium is unique, and Lemma 4 provides the range of B_i's for which that equilibrium continues to exist although not necessarily uniquely. This is an important finding that we will discuss further in the next section where we analyze our continuity refinement.

Lemma 4 is silent with respect to which of the two groups will be mixing and which will not be voting. What it says is that if $B_m < 1$ and $B_n < 1$ it can be either that the m individuals do not vote and the n individuals mix, or that the n individuals do not vote and the m individuals mix. The next lemma shows that for a given pair (B_i, B_j) these two "Pure-Mix" equilibria of Lemma 4 cannot co-exist.[10]

Lemma 5. *For $B_i < 1$ and $B_j < 1$, $B_i \geq \underline{B_i}$ and $B_j \geq \underline{B_j}$ are mutually exclusive, for all $i, j \in \{m, n\}$ and $i \neq j$.*

Proof. Suppose not and consider the $(B_i, B_j)-$space. We first show that $\underline{B_i} > B_j$, or equivalently:

$$\begin{aligned}(j-1)B_j &> (j-1)B_j^{\frac{j}{j-1}} \\ 1 &> B_j^{\frac{1}{j-1}}.\end{aligned}$$

For the same reason we also have $\underline{B_j} > B_i$. Then, $B_i \geq \underline{B_i}$ and $\underline{B_i} > B_j$ imply $B_i > B_j$, while $B_j \geq \underline{B_j}$ and $\underline{B_j} > B_i$ imply $B_j > B_i$ leading to a contradiction. □

3.2 "Mix-Mix" equilibria

Lemmas 1 to 5 completely characterized the "Pure-Pure" and "Pure-Mix" equilibria of our voting game. We are left to analyze the "Mix-Mix" equilibria. Obviously, in any "Mix-Mix" equilibrium the voting conditions (1) for the two groups hold with equality. That is, the best reply under mixing for the m individuals is defined by:

$$A_m = B_m \tag{7}$$

and the best reply under mixing for the n individuals is defined by:

$$A_n = B_n \tag{8}$$

Since we have imposed the condition that all individuals within a group employ the same strategy, it suffices to focus on the mixing condition of a single m and on that of a a single n individual and analyze the intersections between these two in the (p_m, p_n)-space. These intersections are "Mix-Mix" equilibrium pairs (p_m^*, p_n^*), which are what we are after in this Subsection.

In contrast with the "Pure-Pure" and "Pure-Mix" cases, the "Mix-Mix" case entails solving a system of two polynomial equations of arbitrary power –expressions

[10]Meaning that for a given pair (B_i, B_j) we either have m individuals not voting and n mixing *or* n individuals not voting and m mixing (but not both).

(7) and (8)–, and thus there is no general algebraic solution for equilibrium strategies (by the Abel-Ruffini theorem which states that there is no general algebraic solution to polynomial equations of degree five or higher with arbitrary coefficients). Instead, in order to analyze them we will use a number of indirect results about the space these equilibria lie on. For this it is useful to distinguish among the three cases: $B_m = B_n$, $B_m > B_n$, and $B_m < B_n$. Thus, by (7) and (8), these translate into $A_m = A_n$, $A_m > A_n$, and $A_m < A_n$. Analyzing $A_m = A_n$ will greatly help the analysis of the other two cases.

The set of points for which $A_m = A_n$ is depicted by the two black lines of Figure 1; the increasing and the decreasing one. These two black lines divide the (p_m, p_n)-space in four regions depending on the ranking of A_m and A_n. Keep in mind that the two mixing conditions are defined in the same space; dividing the space in these four regions will help us analyze how the intersections of the two mixing conditions behave.[11]

In Appendix A we characterize the set $A_m = A_n$, through a series of lemmas. First, we find the four points where these two lines touch the edges of the (p_m, p_n)-space (Lemma 6). Second, we characterize the decreasing line connecting the top-left corner with the bottom-right corner (Lemma 7). Finally we characterize the increasing line (Lemmas 8 and 9).[12] We summarize this series of lemmas in Proposition 1.

Proposition 1. *The only points* $(p_m, p_n) \in (0,1)^2$ *satisfying condition*

$$A_m = A_n \tag{9}$$

are the points along the line $p_m + p_n = 1$ *and along a continuous line that goes from* $(0,0)$ *to* $(p_m^{***}, 1)$ *where* $p_m^{***} = \frac{n(n-1)}{n(n-1)+\sqrt{n(n-1)(m-n)(m-n+1)}}$.

Proof. See Appendix A. □

The two lines that we have characterized divide the (p_m, p_n)-space in four regions: two where $A_m > A_n$ and two where $A_m < A_n$. Since we are examining "Mix-Mix" equilibria, these two inequalities directly translate into conditions on B_m and B_n that must be satisfied in "Mix-Mix" equilibria (see (7) and (8)). Thus when $B_m > B_n$ all equilibria lie in one of the two regions where $A_m > A_n$, and when $B_m < B_n$ all equilibria lie in one of the two regions where $A_m < A_n$.

What we prove next (see Appendix B) is that the two mixing conditions cross at most once in each of the four regions delimited by the set of points such that $A_m = A_n$ (see Figure 1). This implies that we always have at most two "Mix-Mix" equilibria, one where $p_m^* + p_n^* < 1$ which we name "Mix-Mix 1", and one where $p_m^* + p_n^* > 1$ which we name "Mix-Mix 2". Including the "Pure-Mix" equilibrium (Lemma 4 characterizes it and Lemma 5 proves it is unique), this shows that we have at most three equilibria. Hence, we prove the following:

[11] At this stage it is interesting to compare our analysis with the one of PR. In particular in Section 6 of PR, they discuss "totally quasi-symmetric equilibria", which are what we call "Mix-Mix" equilibria. However they analyze two special cases, which in our notation are: i) $p_m = p_n$ and $m = n$ and ii) $p_m + p_n = 1$. In terms of our Figure 1 it means that they analyze equilibria that might arise along the two diagonals (in the case of the 45-degree line, they also assume $m = n$).

[12] Note that, as it will be explained in more detail later, the fact that the increasing line is in fact increasing is not needed. What is only needed is that it crosses the decreasing line once.

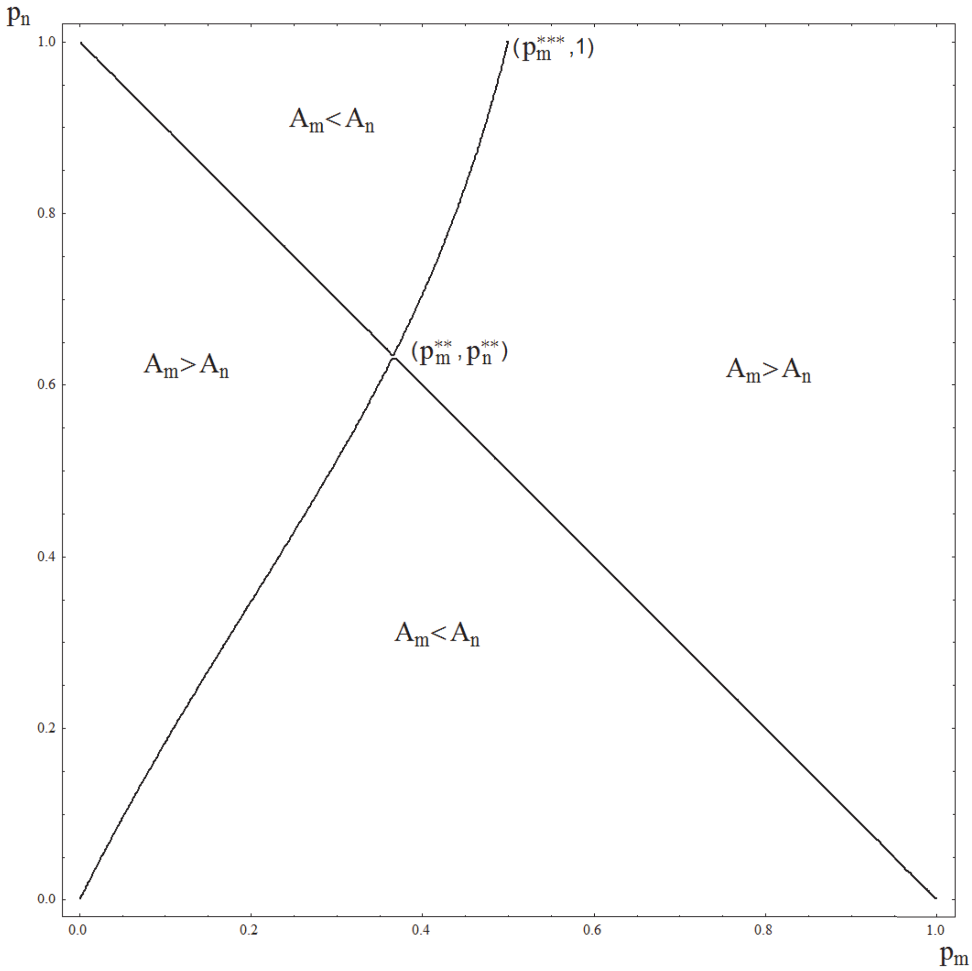

Figure 1: Set of points in the (p_m, p_n)-space according to whether $A_m \gtreqless A_n$ when $m = 3$ and $n = 2$.

Theorem 1. *If $B_i > \underline{B_i}$ for individuals of both groups $i \in \{m, n\}$, there exists a unique QSNE such that $p^*_m = p^*_n = 0$.*

If $B_i > \underline{B_i}$ and $B_j \leq \underline{B_j}$ with $i \in \{m, n\}$ and $i \neq j$ there exists a unique QSNE such that $p^*_i = 0$ and $p^*_j = 1 - B_j^{\frac{1}{j-1}}$.

If $B_i \leq \underline{B_i}$ for all $i \in \{m, n\}$, there are at most three QSNE, one "Pure-Mix" and two "Mix-Mix" ones.

Proof. The first statement follows from Lemma 1. The second statement follows from Lemmas 2 to 5. The third statement is proved in Appendix B. □

Theorem 1 concludes the first part of the paper, establishing that there are at most three equilibria. Returning to PR, the analysis that they carry out for two special cases makes them "[...] conjecture that the class of all totally quasi-symmetric equilibria is much larger than those we have been able to investigate."[13] However, we showed that this class of equilibria – which we call "Mix-Mix"– admits, at most, two equilibria.

What we do in the next Section is propose a refinement and show that this refinement always pins down a unique equilibrium. We first deliver the intuition on how the continuity refinement works by means of a numerical example.

4 Continuous Refinement and Uniqueness

In this section we consider the simplest non-trivial example: $m = 3$ and $n = 2$. We compute and depict the three equilibria and give the intuition how the continuity refinement pins down a unique equilibrium. All qualitative features of this numerical example hold for any m and n.

For the sake of the numerical example, we fix $\Delta\pi_m = \frac{n}{m}\Delta\pi_n$. This parametrization complies with the application we discuss in Section 5 but it also has a simple interpretation: the individual gain of winning of an individual belonging to the big group m, is smaller than the individual gain of an individual of group n. Moreover, how smaller is governed by the ratio of the two group sizes. This implies that we can write $B_n = B$ and $B_m = \frac{m}{n}B$, and thus we have only one parameter B simplifying the exposition greatly.

For any B we examine all the types of equilibria: "Pure-Pure", "Mix-Mix", and "Pure-Mix". We find that for low B there are three different types of equilibria, which are depicted in the first row of Figure 2, while for larger B we only have the equilibrium depicted in the second row. In fact the unique equilibrium for B sufficiently large corresponds to the characterization in Lemmas 1 to 5.

Now, take for example $B = 1/3$ and consider Figure 2. There are three different equilibria we can have: two equilibria where both types of individuals are mixing ("Mix-Mix 1" and "Mix-Mix 2") and one where the n individuals vote for sure and the m individuals mix. This last equilibrium is of the form "Pure-Mix". However, it involves individuals from one group voting with certainty, and thus it is in sharp contrast with the equilibrium characterized in Lemmas 1 to 5, where individuals from one group do not vote. Since this equilibrium will be ruled out by our continuity

[13]PR, page 10.

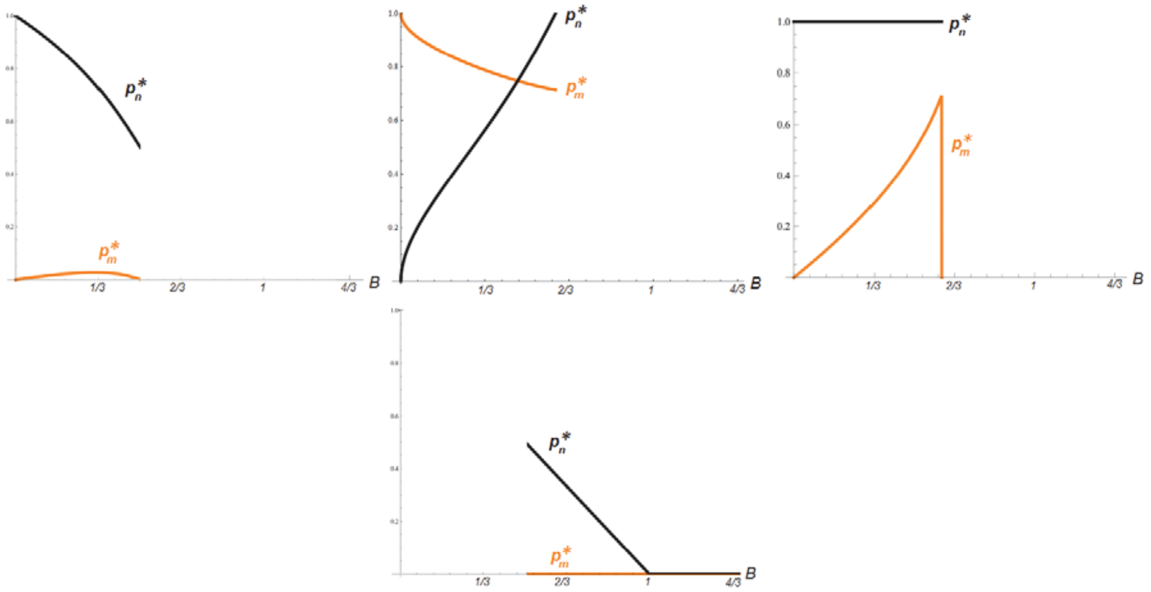

Figure 2: First row, first panel: "Mix-Mix 1". First row, second panel: "Mix-Mix 2". Second row: "High B" equilibrium.

refinement, throughout the rest of the paper we refer to "Pure-Mix" as the one characterized in Lemmas 1 to 5 (with $p_i^* = 0$). These three equilibria correspond to Theorem 1. On the other hand for larger (but not too large) B, say $B \in (2/3, 1)$, there is only one type of equilibrium, the one where the n individuals are mixing, and the m individuals are abstaining for sure (note that in this case $B_m < 1 < B_n$). And naturally, for $B \geq 1$, no-one votes. These two last cases are the "High B" equilibrium in Figure 2 which is composed of the "Pure-Pure" equilibrium of Lemma 1 and the "Pure-Mix" equilibrium of Lemmas 1 to 5.

Starting from sufficiently high B we have to be in the "High B" equilibrium. As we keep decreasing B, at some point we need to switch to one (or more) equilibria from the first row of Figure 2. However the only equilibrium out of the three that involves no jumps in the probabilities of voting, is the "Mix-Mix 1" equilibrium. We show that this continuous equilibrium (depicted in Figure 2) exists and is unique for all (m, n) and for all (B_i, B_j).

The "Mix-Mix 1" and "Mix-Mix 2" equilibria can also be seen in the mixing condition graphs depicted in Figure 4 in the (p_m, p_n)-space. The red (blue) lines represent the mixing condition for a n (m) individual for four values of B, $\{0.166, 0.333, 0.5, 0.61\}$. In particular, we chose the third value to be the minimum B such that the "Mix-Mix 1" equilibrium disappears (bottom-left panel), and the fourth value to be the minimum B such that even the "Mix-Mix 2" equilibrium disappears. The black lines in Figure 4 are the set of (p_m, p_n) satisfying $A_m = A_n$ as in Figure 1.

In each panel, the equilibrium on the left side of the panel is the "Mix-Mix 1" equilibrium. It converges to the "Pure-Mix" equilibrium $(p_m^*, p_n^*) = (0, 0.5)$ as $B \to 0.5$. The equilibrium that lays on the right side of the panel is the "Mix-Mix

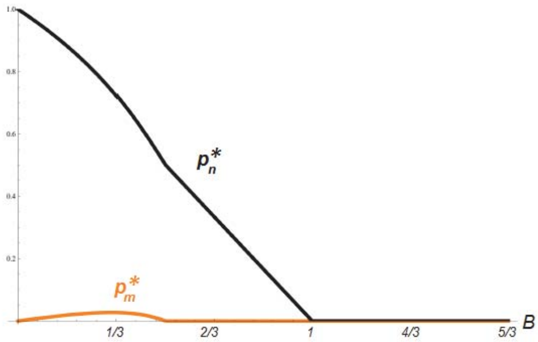

Figure 3: Unique continuous equilibrium

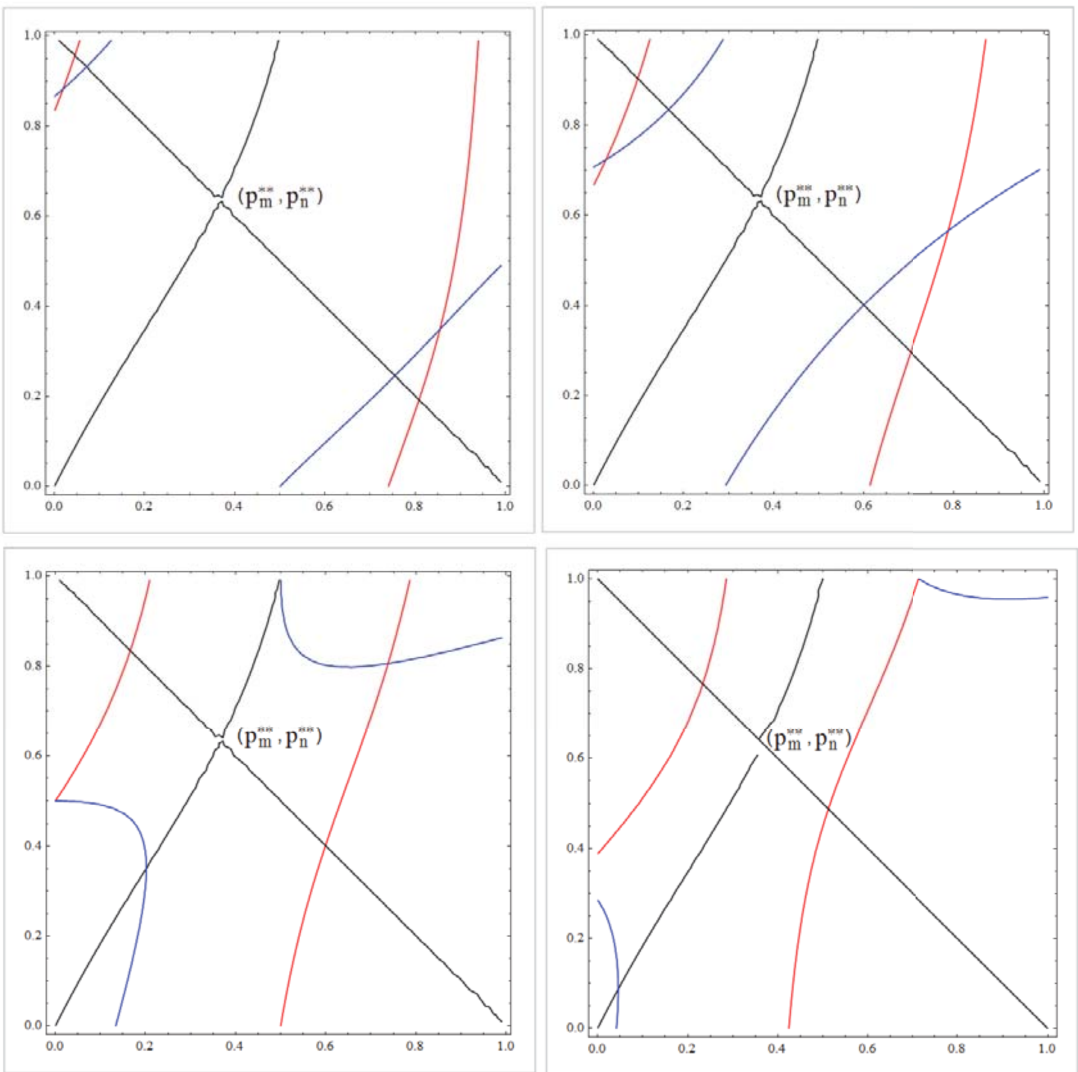

Figure 4: Mixing conditions in the (p_m, p_n)-space respectively for $B = \{0.1666, 0.3333, 0.5, 0.61\}$.

2" equilibrium.

We use a standard definition of continuity (see for example page 943 of Mas-Colell et al. 1995). This definition means that, for all $i, j \in \{m, n\}$, $i \neq j$, there is a single continuous selection of the equilibrium correspondences $p_i^*(B_i, B_j)$ mapping (B_i, B_j) to equilibrium probabilities of voting.

Which of the three equilibria together with the "High B" unique equilibrium satisfies continuity? First, notice that the three equilibria need that $B_m, B_n < 1$. This means that we can focus on the part of the "High B" equilibrium that is defined for $B_m, B_n < 1$ as well. Hence, we need to find among the three equilibria the one for which p_i^* goes to 0 and at the same time, p_j^* is interior.[14] For this to happen we need that $p_m^* + p_n^* < 1$, which contradicts the third equilibrium in Figure 2 because this equilibrium requires that one of the two p_i^*'s is equal to one. Out of the two "Mix-Mix" equilibria, "Mix-Mix 2" contradicts $p_m^* + p_n^* < 1$.

Therefore we are left with only one candidate, namely "Mix-Mix 1", that together with "High B" equilibrium, might satisfy continuity. First notice that Lemma 4 gives us for every $B_j \in (0, 1)$ the lowest B_i for existence of the "Pure-Mix" equilibrium $p_i^* = 0$ and $p_j^* = 1 - B_j^{\frac{1}{j-1}}$. This satisfies the system $A_i \leq B_i$ and $A_j = B_j$, which implies that the system $A_i = B_i$ and $A_j = B_j$ is also satisfied (A_i's are continuous in p_i's).

Thus we proved the following:

Theorem 2. *There exists a unique pair (p_m^*, p_n^*) such that continuity holds at all $(B_m, B_n) \in \mathbb{R}_{++}^2$, and it is composed of "Mix-Mix 1" and "High B".*

5 Application - voting over redistribution of wealth

Voting over wealth redistribution is a neat setting where our model applies: there are more poor individuals than rich, thus redistribution of wealth yields greater harm to a single rich individual than benefit to a single poor individual, thus complying with our setting of asymmetric individual benefits.

There are m poor individuals whose wealth we normalize to 1 and n rich individuals whose wealth we normalize to 2.[15] Individuals are called to vote between two extreme redistribution policies: no redistribution (alternative N) and full redistribution (alternative M). Since the wealth of the poor is less than the average wealth, a poor m individual would always prefer the higher level of redistribution (alternative M) and, at the same time, a rich n individual would always prefer the lower level (alternative N). As a reminder, we assume that $m > n > 1$.

Under full redistribution each individual ends up with the average wealth in the economy, which is:

$$\frac{2n + m}{n + m}$$

[14] In terms of Figure 1 and Figure 4, equilibria must converge to the horizontal or vertical axis. This happens in the third panel of Figure 4.

[15] It will be clear that these wealth level assumptions are qualitatively without loss of generality, since different wealth levels would just re-scale the results in the cost parameter c.

and under no redistribution everyone keeps her original wealth. Thus,

$$\begin{aligned}\Delta\pi_m &= \frac{2n+m}{n+m} - 1 = \frac{n}{n+m}\\ \Delta\pi_n &= 2 - \frac{2n+m}{n+m} = \frac{m}{n+m}\end{aligned}$$

Notice that $\Delta\pi_m = \frac{n}{m}\Delta\pi_n$, as in our parametrization in Section 4.

The policy is decided by majority rule with ties being broken evenly, and for simplicity the cost of voting is assumed to be identical for each individual - that is, $c_m = c_n = c > 0$. This means that $B_m = \frac{c(n+m)}{n}$ and $B_n = \frac{c(n+m)}{m}$. The decision of the individuals is whether to cast a vote for their preferred alternative and pay c, or to not vote at all. Thus, an individual of group i may cast a vote if (1) is met: that is, if $A_m \geq \frac{c(n+m)}{n}$ for a poor individual, and $A_n \geq \frac{c(n+m)}{m}$ for a rich individual. The individuals' payoff is equal to their final wealth minus, possibly, the cost of voting.

From Lemma 4 we know that the "Pure-Mix" $p^*_m = 0$ and $p^*_n = 1 - B_n^{\frac{1}{n-1}}$ exist if and only if $B_m > nB_n - (n-1)B_n^{\frac{n}{n-1}}$. Therefore, if the "Pure-Mix" still exists as c goes to 0 (equivalently, B goes to 0), then "High B" is the unique equilibrium without any need for continuity, and the poor individuals will never vote for any $c > 0$. Notice that the second term of the right-hand side of $B_m \geq nB_n - (n-1)B_n^{\frac{n}{n-1}}$ goes to 0 faster than the other terms in the inequality as B's go to zero, and thus in the limit it is negligible. Then sufficiently close to 0 we are left with only $B_m > nB_n$. By plugging the expressions for B_m and B_n we get:

$$m \geq n^2. \tag{10}$$

This is a necessary and sufficient condition for $p^*_m = 0$ to hold in the unique equilibrium for any $c > 0$. Also, it has a nice interpretation. If the society is sufficiently polarized ($m < n^2$), the poor might vote and redistribution has a chance of winning. However, in a sufficiently non-polarized society ($m \geq n^2$), poors are doomed to lose the election.

How does a change of m affect (p^*_m, p^*_n)? We answer with the support of Figure 5. We fix $n = 3$, and set m so as to initially have a polarized society ($m = 4$, top-left), and gradually decrease the polarization ($m = 5$, top-right, and $m = 6$, bottom-left), until we hit polarization $m = n^2$ ($m = 9$, bottom-right). When we hit this polarization threshold condition (10) is satisfied and the "Mix-Mix 1" equilibrium (which survives continuity) disappears, and we have only "High B" ("Pure-Mix" and "Pure-Pure").

Consider $n = 3$ and $m = 4$. Since the society has (slightly) more poor than rich individuals, the average wealth is closer to the wealth of a poor individual than to the wealth of a rich individual, thus if full redistribution wins, the individual loss of a single rich individual is greater than the individual gain of a single poor individual. For this reason, an n rich individual is willing to vote for greater B's than an m poor individual. In other words, a rich has more at stake than a poor, and thus is willing to face a greater cost of voting. Therefore, p^*_n turns positive for greater B's than p^*_m, as we can see in Figure 5.

Consider an increase of m ($n = 3$ and $m = 5$, or 6). This has the effect of sharpening the asymmetry in willingness to face the cost of voting between rich and

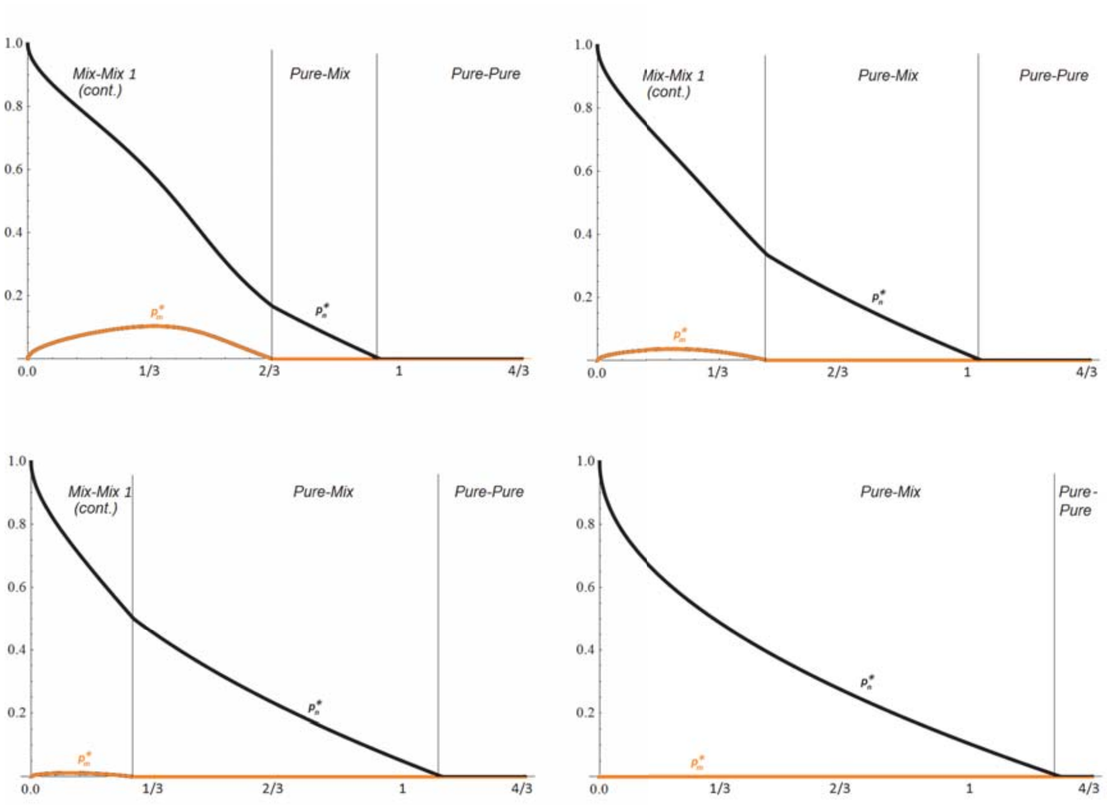

Figure 5: Effects on voting probabilities of increasing m, keeping $n = 3$. First row left panel $m = 4$, right panel $m = 5$. Second row, left panel $m = 6$ right panel $m = 9$.

poor: in fact, now, p_n^* turns positive for even greater B's (the rich has even more at stake to lose in case of full redistribution), while p_m^* turns positive for even lower B's (the poor has even less at stake to win in case of full redistribution). This widens the "Pure-Mix" region (see Figure 5).

If the increase in m reaches the polarization threshold when $m = n^2$ ($n = 3$ and $m = 9$), the poor has so little at stake that she is nowhere willing to face the cost of voting with positive probability in equilibrium.[16] A further increase in m would still imply $p_m^* = 0$ everywhere, and further increases the willingness to vote of the rich (i.e., p_n^* increases for any given B, and p_n^* turns positive for greater B's).

This could be read as a "poverty trap": the greater is the share of poor in a society, the less likely is redistribution of resources to be the outcome of a democratic process (and if $m \geq n^2$ this probability is zero). Thus, the poor might have an incentive to attempt non-democratic channels to exit the policy trap.

[16]Remember that if $m \geq n^2$ the equilibrium is unique without the need for continuity selection.

6 Appendix A

We prove Proposition 1 by way of the following lemmata. See Figure 1.

Lemma 6. *The only points satisfying (9) and* $(p_m, p_n) \in \{0,1\}^2$ *are:* $(0,0)$, $(0,1)$, $(1,0)$, *and* $(p_m^{***}, 1)$, *with* $p_m^{***} = \frac{n(n-1)}{n(n-1)+\sqrt{n(n-1)(m-n)(m-n+1)}}$. *Also,* $p_m^{***} = 1$ *iff* $m = n$.

Proof. By continuity of A_m and A_n in p_m and p_n, in order to analyze the behavior of A_m and A_n in $(p_m, p_n) \in \{0,1\}^2$ we can compute the following limits for A_m

$$\begin{aligned}
\lim_{p_m \to 0} A_m &= np_n(1-p_n)^{n-1} + (1-p_n)^n \\
\lim_{p_m \to 1} A_m &= \begin{cases} p_n^n + np_n^{n-1}(1-p_n) & \text{if } m = n \\ \binom{m-1}{n} p_n^n & \text{if } m = n+1 \\ 0 & \text{if } m > n+1 \end{cases} \\
\lim_{p_n \to 0} A_m &= (1-p_m)^{m-1} \\
\lim_{p_n \to 1} A_m &= \binom{m-1}{n-1} p_m^{n-1}(1-p_m)^{m-n} + \binom{m-1}{n} p_m^n (1-p_m)^{m-n-1}
\end{aligned}$$

and for A_n

$$\begin{aligned}
\lim_{p_m \to 0} A_n &= (1-p_n)^{n-1} \\
\lim_{p_m \to 1} A_n &= \begin{cases} p_n^{n-1} & \text{if } m = n \\ 0 & \text{if } m > n \end{cases} \\
\lim_{p_n \to 0} A_n &= m(1-p_m)^{m-1} + (1-p_m)^m \\
\lim_{p_n \to 1} A_n &= \binom{m}{n} p_m^n (1-p_m)^{m-n} + \binom{m}{n-1} p_m^{n-1}(1-p_m)^{m-n+1}
\end{aligned}$$

From the above,

- if $p_m = 0$, (9) holds iff $p_n = 0$ or $p_n = 1$
- if $p_m = 1$, (9) holds iff $p_n = 0$ or (see Appendix C) $p_n = 1$ and $m = n$
- if $p_n = 0$, (9) holds iff $p_m = 0$ or $p_m = 1$
- if $p_n = 1$, (9) is equivalent to

$$\begin{aligned}
\binom{m-1}{n-1} p_m^{n-1}(1-p_m)^{m-n} + \binom{m-1}{n} p_m^n (1-p_m)^{m-n-1} &= \binom{m}{n} p_m^n (1-p_m)^{m-n} + \binom{m}{n-1} p_m^{n-1}(1-p_m)^{m-n+1} \\
\binom{m-1}{n-1}(1-p_m) + \binom{m-1}{n} p_m &= \binom{m}{n} p_m(1-p_m) + \binom{m}{n-1}(1-p_m)^2 \qquad (11)
\end{aligned}$$

If $m = n$, (11) boils down to

$$1 - p_m = p_m(1-p_m) + m\,(1-p_m)^2$$

whose unique solution is $p_m = 1$.
If $m > n$, (11) boils down to

$$\frac{(1-p_m)}{m-n} + \frac{p_m}{n} = \frac{mp_m(1-p_m)}{n(m-n)} + \frac{m(1-p_m)^2}{(m-n)(m-n+1)}$$

Solving the simple polynomial in the last expression we see that p_m^{***} is indeed one of its two roots (the second root has to be discarded since it is greater than 1). $\square$

Lemma 7. *Equation* $p_m + p_n = 1$ *solves* $A_m = A_n$ $\forall (p_m, p_n) \in [0,1]^2$.

Proof. From (2) plug A_m and A_n into (9), simplify for $(1-p_m)^m(1-p_n)^n$, and use $p_n = 1 - p_m$ to obtain

$$\sum_{s=0}^{n} \binom{m-1}{s}\binom{n}{s} p_m^{s-s}(1-p_m)^{-s-1+s} + \sum_{s=0}^{n-1} \binom{m-1}{s}\binom{n}{s+1} p_m^{s-s-1}(1-p_m)^{-s-1+s+1}$$
$$= \sum_{s=0}^{n-1} \binom{m}{s}\binom{n-1}{s} p_m^{s-s-1}(1-p_m)^{-s+s} + \sum_{s=0}^{n-1} \binom{m}{s+1}\binom{n-1}{s} p_m^{s+1-s-1}(1-p_m)^{-s-1+s}$$

$$p_m \sum_{s=0}^{n} \binom{m-1}{s}\binom{n}{s} + (1-p_m) \sum_{s=0}^{n-1} \binom{m-1}{s}\binom{n}{s+1}$$
$$= (1-p_m) \sum_{s=0}^{n-1} \binom{m}{s}\binom{n-1}{s} + p_m \sum_{s=0}^{n-1} \binom{m}{s+1}\binom{n-1}{s}$$

$$p_m \sum_{s=0}^{n} \binom{m-1}{s}\binom{n}{n-s} + (1-p_m) \sum_{s=0}^{n-1} \binom{m-1}{s}\binom{n}{n-s-1}$$
$$= (1-p_m) \sum_{s=0}^{n-1} \binom{m}{s}\binom{n-1}{n-s-1} + p_m \sum_{s=0}^{n-1} \binom{m}{s+1}\binom{n-1}{n-s-1}$$

$$p_m \binom{m+n-1}{n} + (1-p_m)\binom{m+n-1}{n-1} = (1-p_m)\binom{m+n-1}{n-1} + p_m \binom{m+n-1}{n}$$
$$0 = 0$$

where in the second-to-last step we used the symmetry rule for binomial coefficients, and in the last step we used Vandermonde's identity.[17] $\square$

Next we characterize the set of points $A_m = A_n$ that are depicted by an increasing line in the (p_m, p_n)-space by means of two lemmas. In Lemma 8 We show that there exists a point (p_m^{**}, p_n^{**}) along the decreasing line which divides the neighborhoods of the decreasing line into two parts:

1. The first part is the one connecting (p_m^{**}, p_n^{**}) and $(1,0)$, where we prove that increasing p_m (i.e., moving to the right of the line), increases A_m *more* than A_n. Since exactly along the line $A_m = A_n$, this result implies that to the right of the segment connecting (p_m^{**}, p_n^{**}) and $(1,0)$ we have $A_m > A_n$ and to its left we have $A_m < A_n$.

[17] Vandermonde's identity states that $\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}$ for $m, n, r \in \mathbb{N}_0$.

2. The second part is the one connecting $(0,1)$ and (p_m^{**}, p_n^{**}), where we prove that increasing p_m (i.e., moving to the right of the line), increases A_m *less* than A_n. Since exactly along the line, $A_m = A_n$, this result implies that to the right of the segment connecting $(0,1)$ and (p_m^{**}, p_n^{**}) we have that $A_m < A_n$ and to its left $A_m > A_n$.

Lemma 8. *$\exists!(p_m^{**}, p_n^{**}) \in (0,1)^2$ with $p_m^{**} + p_n^{**} = 1$ such that*

$$\left.\frac{\partial A_m}{\partial p_m}\right|_{p_m+p_n=1} > \left.\frac{\partial A_n}{\partial p_m}\right|_{p_m+p_n=1} \quad \textit{iff } p_m > p_m^{**} \textit{ (or equivalently } p_n < p_n^{**}\textit{)}$$

Also, if $m = n$, then $p_m^{**} = p_n^{**} = \frac{1}{2}$, and if $m > n$, then $p_m^{**} \in (0, \frac{1}{2})$ and $p_n^{**} \in (\frac{1}{2}, 1)$.

In particular,

$$p_m^{**} = \frac{n(n-1)}{n(n-1) + \sqrt{m(m-1)n(n-1)}} \quad \textit{and } p_n^{**} = 1 - p_m^{**}$$

Proof. For notation simplicity and for the sake of space we define the following

$$\begin{aligned} \tilde{p}_{s,m} &= \left(\frac{p_m}{1-p_m}\right)^s \\ \tilde{p}_{s,n} &= \left(\frac{p_n}{1-p_n}\right)^s \end{aligned}$$

Then,

$$\left.\frac{\partial A_m}{\partial p_m}\right|_{p_m+p_n=1} > \left.\frac{\partial A_n}{\partial p_m}\right|_{p_m+p_n=1}$$

$$\left.\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\frac{s+p_m}{p_m(1-p_m)^2} + \sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}\tilde{p}_{s,m}\tilde{p}_{s,n}\frac{s+p_m}{p_m(1-p_m)^2}\frac{p_n}{1-p_n}\right|_{p_m+p_n=1} >$$
$$\left.\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\frac{s}{p_m(1-p_m)(1-p_n)} + \sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\frac{s+1}{(1-p_m)^2(1-p_n)}\right|_{p_m+p_n=1}$$

by noticing that $p_m + p_n = 1$ implies $\tilde{p}_{s,m}\tilde{p}_{s,n} = 1$ the above inequality simplifies to

$$\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}\frac{s+p_m}{p_m(1-p_m)^2} + \sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}\frac{s+p_m}{p_m^2(1-p_m)} >$$
$$\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}\frac{s}{p_m^2(1-p_m)} + \sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}\frac{s+1}{p_m(1-p_m)^2}$$

$$\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}p_m(s+p_m) + \sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}(1-p_m)(s+p_m) >$$
$$\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}(1-p_m)s + \sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}p_m(s+1)$$

Note that some summands in the above inequality contain s only in the binomial coefficients. By applying to these terms the same procedure at the end of Lemma 7 (i.e. symmetry rule for binomial coefficients and Vandermonde's identity), we get

$$p_m \sum_{s=0}^{n} \binom{m-1}{s}\binom{n}{s} s + p_m^2 \binom{m+n-1}{n} + (1-p_m) \sum_{s=0}^{n-1} \binom{m-1}{s}\binom{n}{s+1} s + p_m(1-p_m)\binom{m+n-1}{n-1} >$$
$$(1-p_m) \sum_{s=0}^{n-1} \binom{m}{s}\binom{n-1}{s} s + p_m \sum_{s=0}^{n-1} \binom{m}{s+1}\binom{n-1}{s} s + p_m \binom{m+n-1}{n}$$

$$p_m \sum_{s=0}^{n} \binom{m-1}{s}\binom{n}{s} s + (1-p_m) \sum_{s=0}^{n-1} \binom{m-1}{s}\binom{n}{s+1} s >$$
$$(1-p_m) \sum_{s=0}^{n-1} \binom{m}{s}\binom{n-1}{s} s + p_m \sum_{s=0}^{n-1} \binom{m}{s+1}\binom{n-1}{s} s + p_m(1-p_m)\left[\binom{m+n-1}{n} - \binom{m+n-1}{n-1}\right]$$

we now analyze the summations left (containing s not only in the binomial coefficient), and use the fact that $\sum_{s=0}^{n} \binom{m}{s}\binom{n}{s} s = n\binom{m+n-1}{n}$ and that $\sum_{s=0}^{n} \binom{m}{s}\binom{n}{s+1} s = m\binom{m+n-1}{n-2}$,[18] and get

$$np_m \binom{m+n-2}{n} + (m-1)(1-p_m)\binom{m+n-2}{n-2} >$$
$$(n-1)(1-p_m)\binom{m+n-2}{n-1} + (n-1)p_m\binom{m+n-2}{n} + p_m(1-p_m)\left[\binom{m-n-1}{n} - \binom{m-n-1}{n-1}\right]$$

$$(m-1)(1-p_m)\binom{m+n-2}{n-2} >$$
$$(n-1)(1-p_m)\binom{m+n-2}{n-1} - p_m\binom{m+n-2}{n} + p_m(1-p_m)\left[\binom{m-n-1}{n} - \binom{m-n-1}{n-1}\right]$$

and simplifying by $\frac{(m+n-2)!}{(m-2)!(n-2)!}$ we get

$$\begin{aligned}
\frac{1-p_m}{m} &> \frac{1-p_m}{m-1} - \frac{p_m}{n(n-1)} + p_m(1-p_m)\frac{(m-n)(m+n-1)}{m(m-1)n(n-1)} \\
-n(n-1)(1-p_m) &> -m(m-1)p_m + p_m(1-p_m)(m-n)(m+n-1) \\
(m-n)(m+n-1)p_m^2 + 2n(n-1)p_m - n(n-1) &> 0 \\
p_m &> \frac{n(n-1)}{n(n-1)+\sqrt{m(m-1)n(n-1)}} = p_m^{**}
\end{aligned}$$

18

$$\begin{aligned}
\sum_{s=0}^{n} \binom{m}{s}\binom{n}{s} s &= \sum_{s=0}^{n} \binom{m}{s}\frac{n!}{s!(n-s)!} s \\
&= \sum_{s=0}^{n} \binom{m}{s}\frac{n!}{(s-1)!(n-s)!} = \sum_{s=0}^{n} \binom{m}{s}\frac{n!}{(s-1)!(n-1-s+1)!} \\
&= \sum_{s=0}^{n} \binom{m}{s}\frac{(n-1)!}{(s-1)!((n-1)-(s-1))!} n = n\binom{m+n-1}{n}
\end{aligned}$$

Where the last equality follows from Valdemore's identity. The calculations for the other summation are similar.

If $m = n$ it is trivial to see that $p_m^{**} = \frac{1}{2}$. But notice also that p_m^{**} decreases in m, and hence by $m > n$, $p_m^{**} \in (0, \frac{1}{2})$ and $p_n^{**} \in (\frac{1}{2}, 1)$. $\square$

The next Lemma proves that for each p_i there exist at most two p_j that satisfy $A_i = A_j$, for $i, j \in \{m, n\}$, $i \neq j$.

Lemma 9. *$\forall (m, n) \in \mathbb{Z}^2$ with $m \geq n \geq 2$, $\forall p_n \in (0, 1)$ there exist at most two values of $p_m \in (0, 1)$ such that $A_m = A_n$.*

Proof. Taking p_n as given, $A_m = A_n$ is a polynomial in p_m.[19] In order to prove that there are at most two roots of p_m that solve $A_m = A_n$ we will invoke Descartes' Rule of Signs[20] to show that the coefficients of $(p_m^0, p_m^1, p_m^2, ...)$ change sign at most twice. Descartes' Rule bounds the number of roots in the $x \in (0, +\infty)$ interval of a polynomial in x. Using Jacobi's substitution[21] we rewrite the polynomial in terms of $y = \frac{p_m}{1-p_m}$, so that $p_m \in (0, 1)$ implies $y \in (0, +\infty)$. Since y is a one-to-one mapping $(0, 1) \to (0, +\infty)$ bounding the number of roots of $A_m = A_n$ in $y \in (0, +\infty)$ implies bounding the number of roots of $A_m = A_n$ in $p_m \in (0, 1)$.

Simplify $(1 - p_m)^m (1 - p_n)^n$ in $A_m = A_n$ and get:

$$\sum_{s=0}^{n} \binom{m-1}{s} \binom{n}{s} p_m^s (1-p_m)^{-s-1} p_n^s (1-p_n)^{-s} + \sum_{s=0}^{n-1} \binom{m-1}{s} \binom{n}{s+1} p_m^s (1-p_m)^{-s-1} p_n^{s+1} (1-p_n)^{-s-1} =$$
$$= \sum_{s=0}^{n-1} \binom{m}{s} \binom{n-1}{s} p_m^s (1-p_m)^{-s} p_n^s (1-p_n)^{-s-1} + \sum_{s=0}^{n-1} \binom{m}{s+1} \binom{n-1}{s} p_m^{s+1} (1-p_m)^{-s-1} p_n^s (1-p_n)^{-s-1}$$

We apply Jacobi's substitution setting $\alpha = 0$ and $\beta = 1$ to look for the roots in $(0, 1)$:

$$y = \frac{p_m}{1 - p_m} \tag{12}$$

which implies:

$$p_m = \frac{y}{y+1} \tag{13}$$

We plug 13 into equation $A_m = A_n$:

$$\sum_{s=0}^{n} \binom{m-1}{s} \binom{n}{s} y^s (1-p_m)^{-1} p_n^s (1-p_n)^{-s} + \sum_{s=0}^{n-1} \binom{m-1}{s} \binom{n}{s+1} y^s (1-p_m)^{-1} p_n^{s+1} (1-p_n)^{-s-1} =$$
$$= \sum_{s=0}^{n-1} \binom{m}{s} \binom{n-1}{s} y^s p_n^s (1-p_n)^{-s-1} + \sum_{s=0}^{n-1} \binom{m}{s+1} \binom{n-1}{s} y^{s+1} p_n^s (1-p_n)^{-s-1} \tag{14}$$

Note that using (13) we can get rid of terms $(1 - p_m)^{-1}$:

$$\begin{aligned} (1-p_m)^{-1} &= \left(1 - \frac{y}{y+1}\right)^{-1} \\ &= \left(\frac{1}{y+1}\right)^{-1} \\ &= y + 1 \end{aligned}$$

[19] This is of course without loss of generality, as we can take p_m as given making $A_m = A_n$ a polynomial in p_n.

[20] Page 28, Corollary 1, Prasolov (2001).

[21] Page 28, Prasolov (2001).

Therefore (6) becomes

$$\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}y^s(1+y)p_n^s(1-p_n)^{-s}+\sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}y^s(1+y)p_n^{s+1}(1-p_n)^{-s-1}=$$
$$=\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}y^s p_n^s(1-p_n)^{-s-1}+\sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}y^{s+1}p_n^s(1-p_n)^{-s-1} \tag{15}$$

In (15) we can add the n^{th} element to all summations which go from 0 to $n-1$, because we are adding a term which equals 0. This procedure will allow us to put together all summations.

$$\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}y^s(1+y)p_n^s(1-p_n)^{-s}+\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s+1}y^s(1+y)p_n^{s+1}(1-p_n)^{-s-1}$$
$$=\sum_{s=0}^{n}\binom{m}{s}\binom{n-1}{s}y^s p_n^s(1-p_n)^{-s-1}+\sum_{s=0}^{n}\binom{m}{s+1}\binom{n-1}{s}y^{s+1}p_n^s(1-p_n)^{-s-1}$$

Now we have to deal with the fact that the two summations of the left-hand side have y both to the power of s and $s+1$, so that we separate these two powers, and get:

$$\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}y^s p_n^s(1-p_n)^{-s}+\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}y^{s+1}p_n^s(1-p_n)^{-s}+\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s+1}y^s p_n^{s+1}(1-p_n)^{-s-1}$$
$$+\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s+1}y^{s+1}p_n^{s+1}(1-p_n)^{-s-1}$$
$$=\sum_{s=0}^{n}\binom{m}{s}\binom{n-1}{s}y^s p_n^s(1-p_n)^{-s-1}+\sum_{s=0}^{n}\binom{m}{s+1}\binom{n-1}{s}y^{s+1}p_n^s(1-p_n)^{-s-1}$$

Now we can simply put together all summations with y^s and all summations with y^{s+1}, and get:

$$\sum_{s=0}^{n}y^s\left[\binom{m-1}{s}\binom{n}{s}p_n^s(1-p_n)^{-s}+\binom{m-1}{s}\binom{n}{s+1}p_n^{s+1}(1-p_n)^{-s-1}-\binom{m}{s}\binom{n-1}{s}p_n^s(1-p_n)^{-s-1}\right]+$$
$$+\sum_{s=0}^{n}y^{s+1}\left[\binom{m-1}{s}\binom{n}{s}p_n^s(1-p_n)^{-s}+\binom{m-1}{s}\binom{n}{s+1}p_n^{s+1}(1-p_n)^{-s-1}-\binom{m}{s+1}\binom{n-1}{s}p_n^s(1-p_n)^{-s-1}\right]=0$$

We add a 0 element to both summations ($s=n+1$ in the first, and $s=-1$ in the second):

$$\sum_{s=0}^{n+1}y^s\left[\binom{m-1}{s}\binom{n}{s}p_n^s(1-p_n)^{-s}+\binom{m-1}{s}\binom{n}{s+1}p_n^{s+1}(1-p_n)^{-s-1}-\binom{m}{s}\binom{n-1}{s}p_n^s(1-p_n)^{-s-1}\right]+$$
$$+\sum_{s=-1}^{n}y^{s+1}\left[\binom{m-1}{s}\binom{n}{s}p_n^s(1-p_n)^{-s}+\binom{m-1}{s}\binom{n}{s+1}p_n^{s+1}(1-p_n)^{-s-1}-\binom{m}{s+1}\binom{n-1}{s}p_n^s(1-p_n)^{-s-1}\right]=0$$

In order to put together these two summations we make a variable change in the first summation: $s = q + 1$ to get:

$$\sum_{q=-1}^{n} y^{q+1} \left[\binom{m-1}{q+1}\binom{n}{q+1} p_n^{q+1}(1-p_n)^{-q-1} + \binom{m-1}{q+1}\binom{n}{q+2} p_n^{q+2}(1-p_n)^{-q-2} - \binom{m}{q+1}\binom{n-1}{q+1} p_n^{q+1}(1-p_n)^{-q-2} + \sum_{s=-1}^{n} y^{s+1} \left[\binom{m-1}{s}\binom{n}{s} p_n^{s}(1-p_n)^{-s} + \binom{m-1}{s}\binom{n}{s+1} p_n^{s+1}(1-p_n)^{-s-1} - \binom{m}{s+1}\binom{n-1}{s} p_n^{s}(1-p_n)^{-s-1} = 0 \right.\right.$$

We can finally put together the two summations, and clearly set $s = q$. At the same time we simplify $\frac{(m-1)!(n-1)!}{s!s!(m-s-2)!(n-s-2)!}$, and get:

$$\sum_{s=-1}^{n} y^{s+1} \left[\frac{n}{(s+1)^2(n-s-1)} p_n^{s+1}(1-p_n)^{-s-1} + \frac{n}{(s+1)^2(s+2)} p_n^{s+2}(1-p_n)^{-s-2} - \frac{m}{(s+1)^2(m-s-1)} p_n^{s+1}(1-p_n)^{-s-2} + \frac{n}{(m-s-1)(n-s)(n-s-1)} p_n^{s}(1-p_n)^{-s} + \frac{n}{(m-s-1)(s+1)(n-s-1)} p_n^{s+1}(1-p_n)^{-s-1} - \frac{m}{(m-s-1)(s+1)(n-s-1)} p_n^{s}(1-p_n)^{-s-1} = 0 \right.$$

Now collect $p_n^s(1-p_n)^{-s-2}$ and get:

$$\sum_{s=-1}^{n} y^{s+1} p_n^s (1-p_n)^{-s-2} \left[\frac{n}{(s+1)^2(n-s-1)} p_n(1-p_n) + \frac{n}{(s+1)^2(s+2)} p_n^2 - \frac{m}{(s+1)^2(m-s-1)} p_n + \frac{n}{(m-s-1)(n-s)(n-s-1)} (1-p_n)^2 + \frac{n}{(m-s-1)(s+1)(n-s-1)} p_n(1-p_n) - \frac{m}{(m-s-1)(s+1)(n-s-1)} (1-p_n) = 0. \right. \tag{16}$$

We are ready to apply Descartes's rule of signs to the polynomial (6). Let us analyze the sign of the coefficients of y^{s+1}. The term $p_n^s(1-p_n)^{-s-2}$ is always positive[22]. This means that we only care about the number of sign changes of the term in the square bracket, as s goes from -1 to $n-1$.[23] Note that in this interval the term in the square bracket is continuous in s. We need to show is has at most two roots in the interval $s \in (-1, n-1)$ in order to conclude the proof. We multiply the term in the square brackets by $(s+1)^2(s+2)(m-s-1)(n-s-1)(n-s)$ and get:

$$n(s+2)(m-s-1)(n-s)p_n(1-p_n) + n(m-s-1)(n-s-1)(n-s)p_n^2 - m(s+2)(n-s-1)(n-s)p_n + n(s+1)^2(s+2)(1-p_n)^2 + n(s+1)(s+2)(n-s)p_n(1-p_n) - m(s+1)(s+2)(n-s)(1-p_n) = 0 \tag{17}$$

This is a polynomial of degree *three* in s, thus it has potentially three real roots in s, however, to conclude the proof, we need to show that it has at most two roots

[22] Descartes' rule ignores 0-coefficients, so this is true also when $p_n = \{0, 1\}$

[23] As $s \to n$, the coefficient of y^{n+1} is always positive.

in $s \in (-1, n-1)$. Once again we use Jacobi's substitution to look for roots in $s \in (-1, n-1)$. There fore we set $\alpha = -1$ and $\beta = n-1$ and we have:

$$x = \frac{s+1}{n-s-1} \tag{18}$$

or equivalently,

$$s = \frac{x(n-1)-1}{x+1}$$

Since in each term of (6) there are repeating factors, it will be convenient to first substitute $s = \frac{x(n-1)-1}{x+1}$ to each factor separately. Then we have:

$$\begin{aligned}
s+1 &= \frac{x(n-1)-1}{x+1} + \frac{x+1}{x+1} = \frac{xn}{x+1} \\
s+2 &= \frac{x(n-1)-1}{x+1} + 2\frac{x+1}{x+1} = \frac{x(n+1)+1}{x+1} \\
n-s &= n\frac{x+1}{x+1} - \frac{x(n-1)-1}{x+1} = \frac{x+n+1}{x+1} \\
n-s-1 &= n\frac{x+1}{x+1} - \frac{x(n-1)-1}{x+1} - \frac{x+1}{x+1} = \frac{n}{x+1} \\
m-s-1 &= m\frac{x+1}{x+1} - \frac{x(n-1)-1}{x+1} - \frac{x+1}{x+1} = \frac{x(m-n)+m}{x+1}
\end{aligned}$$

Then we plug these factors in each term of (6):

$$\begin{aligned}
n(s+2)(m-s-1)(n-s)p_n(1-p_n) &= \frac{p_n(1-p_n)n}{(x+1)^3}(x(n+1)+1)(x(m-n)+m)(x+n+1) \\
&= \frac{p_n(1-p_n)n}{(x+1)^3}(x^3(n+1)(m-n) + x^2 m(n+1) + x^2(m-n) + xm \\
&\quad + x^2(n+1)^2(m-n) + xm(n+1)^2 + x(m-n)(n+1) + m(n+1))
\end{aligned}$$

$$\begin{aligned}
n(m-s-1)(n-s-1)(n-s)p_n^2 &= \frac{p_n^2 n^2}{(x+1)^3}(x(m-n)+m)(x+n+1) \\
&= \frac{p_n^2 n^2}{(x+1)^3}(x^2(m-n) + x(m-n)(n+1) + mx + (n+1)m)
\end{aligned}$$

$$\begin{aligned}
-m(s+2)(n-s-1)(n-s)p_n &= -\frac{p_n nm}{(x+1)^3}(x(n+1)+1)(x+n+1) \\
&= -\frac{p_n nm}{(x+1)^3}(x^2(n+1) + x(n+1)^2 + x + n + 1)
\end{aligned}$$

$$n(s+1)^2(s+2)(1-p_n)^2 = \frac{(1-p_n)^2 n^3}{(x+1)^3}(x^3(n+1) + x^2)$$

$$
\begin{aligned}
n(s+1)(s+2)(n-s)p_n(1-p_n) &= \frac{p_n(1-p_n)n}{(x+1)^3}(x^2n(n+1)+xn)(x+n+1)\\
&=\frac{p_n(1-p_n)n}{(x+1)^3}(x^3n(n+1)+x^2n(n+1)^2+x^2n+xn(n+1))
\end{aligned}
$$

$$
\begin{aligned}
-m(s+1)(s+2)(n-s)(1-p_n) &= -\frac{(1-p_n)m}{(x+1)^3}(x^2n(n+1)+xn)(x+n+1)\\
&=-\frac{(1-p_n)m}{(x+1)^3}(x^3n(n+1)+x^2n(n+1)^2+x^2n+xn(n+1))
\end{aligned}
$$

All terms are divided by $(x+1)^3$ so we can drop it in (6). We now proceed by first collecting all terms that are multiplied by x^3.

$$
\begin{gathered}
x^3[p_n(1-p_n)n(n+1)(m-n)-(1-p_n)mn(n+1)+n^3(1-p_n)^2(n+1)+n^2(n+1)p_n(1-p_n)] =\\
x^3n(n+1)(1-p_n)(p_n(m-n)-m+n^2(1-p_n)+np_n) =\\
x^3n(n+1)(1-p_n)(p_nm-m+n^2(1-p_n)) =\\
x^3n(n+1)(1-p_n)^2(n^2-m)
\end{gathered}
$$

Next, we collect all terms that are multiplied by x^2.

$$
\begin{gathered}
x^2\{p_n(1-p_n)nm(n+1)+p_n(1-p_n)n(m-n)+p_n(1-p_n)n(n+1)^2(m-n)+n^2p_n^2(m-n)-p_nmn(n+1)\\
+n^3(1-p_n)^2+np_n(1-p_n)n(n+1)^2+n^2p_n(1-p_n)-m(1-p_n)n(n+1)^2-m(1-p_n)n\} =\\
x^2n\{m[(n+1)p_n(1-p_n)+p_n(1-p_n)+(n^2+2n+1)p_n(1-p_n)+np_n^2-p_n(n+1)-(n^2+2n+1)(1-p_n)\\
-(1-p_n)]-np_n(1-p_n)-(n+1)^2np_n(1-p_n)-n^2p_n^2+(n+1)^2np_n(1-p_n)+n^2(1-p_n^2)+np_n(1-p_n)\} =\\
x^2n\{n^2(1-2p_n^2)+m(4p_n-3p_n^2+2n^2p_n+4np_n-n^2p_n^2-2np_n^2-n^2-2n-2)\} =\\
x^2n\{n^2(1-2p_n^2)+m[4p_n-3p_n^2-2-(2n+n^2)(1-2p_n+p_n^2)]\} =\\
x^2n\{n^2(1-2p_n^2)-m(-4p_n+3p_n^2+2+(2+n)n(1-p_n)^2)\}
\end{gathered}
$$

Next, we collect all terms that are multiplied by x.

$$
\begin{gathered}
x\{nmp_n(1-p_n)+m(n+1)^2np_n(1-p_n)+(m-n)(n+1)np_n(1-p_n)+n^2p_n^2(m-n)(n+1)\\
+n^2p_n^2m-p_nmn(n+1)^2-p_nmn+np_n(1-p_n)n(n+1)-m(1-p_n)n(n+1)\} =\\
xn\{m[p_n-p_n^2+(n+1)^2(p_n-p_n^2)+(n+1)(p_n-p_n^2)+n^2p_n^2+np_n^2+np_n^2-p_n(n+1)^2\\
-p_n-(1-p_n)(n+1)]-n^2p_n^2(n+1)\} =\\
xn\{m[p_n-p_n^2-(n^2+2n+1)p_n^2+np_n+p_n-np_n^2-p_n^2+n^2p_n^2+2np_n^2\\
-p_n-(n+1)+p_n(n+1)]-n^2p_n^2(n+1)\} =\\
xn\{m[-3p_n^2-np_n^2+2p_n+2np_n-(n+1)]-n^2p_n^2(n+1)\} =\\
-xn\{m[(3+n)p_n^2-2p_n(1+n)+n+1]+(n+1)n^2p_n^2\}
\end{gathered}
$$

And finally we collect the constant term.

$$
\begin{gathered}
nmp_n(1-p_n)(1+n)+n^2p_n^2(n+1)m-p_nmn(n+1) =\\
nmp_n(n+1)(1-p_n+np_n-1) =\\
nm(n+1)p_n^2(n-1) =\\
nm(n^2-1)p_n^2
\end{gathered}
$$

Adding up the four terms and setting them equal to zero we get the following expression which is a polynomial of degree three in x, namely:

$$\begin{gathered} n(n+1)(1-p_n)^2(n^2-m)x^3+ \\ +n\{n^2(1-2p_n^2)-m(-4p_n+3p_n^2+2+(2+n)n(1-p_n)^2)\}x^2+ \\ -n\{m[(3+n)p_n^2-2p_n(1+n)+n+1]+(n+1)n^2p_n^2\}x+ \\ +mn(n^2-1)p_n^2=0 \end{gathered}$$

Both the intercept[24] and the coefficient of x^3 are positive,[25] which implies that when $x \to -\infty$, the polynomial tends to $-\infty$. These two facts imply that at least one root is negative, and this the above polynomial crosses the positive axis at most twice, and thus it has at most two positive roots in x (and therefore in s). □

Lemma 10 concludes the characterization of the increasing line.

Lemma 10. *There exists a unique and continuous line in the* (p_m, p_n)*-space which satisfies* $A_m = A_n$ *and connects* $(0,0)$ *and* $(p_m^{***}, 1)$ *Furthermore, this line crosses the* $p_m + p_n = 1$ *line once at* (p_m^{**}, p_n^{**}).

Proof. Lemma 7 establishes that the decreasing line connects two out of the four points satisfying $A_m = A_n$ along the edges. The line connecting the remaining two points is continuous and by Lemma 8 crosses the decreasing line once, at (p_m^{**}, p_n^{**}). □

7 Appendix B

The goal of this Appendix is to show that the two mixing conditions cross at most once in each of the four regions delimited by the set of points such that $A_m = A_n$ (see Figure 1). This implies that there are at most two "Mix-Mix" equilibria.

In Figure 4 we depict these mixing conditions of both types of individuals for different B's, and for $m = 3$ and $n = 2$. Note that i's mixing condition might cross the $p_m + p_n = 1$ line either twice, once or zero times. If they cross twice, one crossing will be above and to the left of the point (p_m^{**}, p_n^{**}), and the other will be below and to the right of it. If they only cross once the crossing coincides with point (p_m^{**}, p_n^{**}). This is what we prove in Proposition 2. Moreover, in Proposition 3 we show that the mixing conditions of the n-individual are steeper (flatter) in all points satisfying $A_m \geq A_n$ ($A_m \leq A_n$). Thus the mixing conditions cross once in each of the two regions of Figure 1 where $A_m > A_n$ ($A_m < A_n$) for sufficiently low B's. This, together with the fact that continuity imposes $p_m + p_n < 1$ yields the result.

Proposition 2. *Define* $\hat{B}_i = \max\limits_{p_m+p_n=1} A_i$. *The number of intersections between* $p_m + p_n = 1$ *and i's mixing condition* $A_i = B_i$ *are:*

1. *two, if* $B_i < \hat{B}_i$. *In particular, one with* $p_m < p_m^{**}$ *and one with* $p_m > p_m^{**}$

[24] If $x = 0$, it equals $mn(n^2-1)p_n^2$ which is positive.

[25] $m < n^2$ for the mix-mix equilibrium to exist.

2. *one, if $B_i = \hat{B}_i$. In particular, the one solving $p_m = p_m^{**}$ and $p_n = 1 - p_m^{**}$*

3. *zero, if $B_i > \hat{B}_i$*

   $\forall i \in \{m, n\}$.

Proof. Consider Figure 6. On the vertical axis there is A_i *conditional* on being along the $p_m + p_n = 1$ line, and on the horizontal axis there is p_m. We are going to show that A_i conditional on being along $p_m + p_n = 1$ is increasing in p_m if and only if $p_m < p_m^{**}$. Therefore there exists a $\hat{B}_i = \max_{p_m+p_n=1} A_i$ such that the "Mix-Mix" condition $A_i = B_i$ under $p_m + p_n = 1$ has two, one or zero solutions according to whether $B_i < \hat{B}_i$, $B_i = \hat{B}_i$ or $B_i > \hat{B}_i$.

We use the same manipulations of A_m and A_n used in the Proof of Lemma 7, and hence (7) with $p_n = 1 - p_m$ reads

$$(1-p_m)^{m-1} p_m^{n-1} \left[p_m \binom{m+n-1}{n} + (1-p_m) \binom{m+n-1}{n-1} \right] = B_m \tag{19}$$

and (8) with $p_n = 1 - p_m$ reads

$$(1-p_m)^{m-1} p_m^{n-1} \left[p_m \binom{m+n-1}{n} + (1-p_m) \binom{m+n-1}{n-1} \right] = B_n \tag{20}$$

and notice that the left-hand sides are clearly identical by construction because we know that along $p_m + p_n = 1$ line we have that $A_m = A_n$ (see Lemma 7), whereas the right-hand sides could be unequal. We now analyze the left-hand side, which can be rewritten as:

$$\binom{m+n-1}{n} (1-p_m)^{m-1} p_m^{n-1} \left[p_m + (1-p_m) \frac{n}{m} \right]$$

or, equivalently,

$$\binom{m+n-1}{n} (1-p_m)^{m} p_m^{n} \left[\frac{1}{1-p_m} + \frac{n}{m p_m} \right]$$

which takes value 0 if $p_m \in \{0, 1\}$. Also it increases in p_m if and only if $p_m < p_m^{**}$, because:

$$\begin{aligned}
\frac{\partial}{\partial p_m} \left[\binom{m+n-1}{n} (1-p_m)^{m} p_m^{n} \left[\frac{1}{1-p_m} + \frac{n}{m p_m} \right] \right] &> 0 \\
\frac{\partial}{\partial p_m} \left[m(1-p_m)^{m-1} p_m^{n} + n(1-p_m)^{m} p_m^{n-1} \right] &> 0 \\
m \left[-(m-1)(1-p_m)^{m-2} p_m^{n} + n(1-p_m)^{m-1} p_m^{n-1} \right] + n \left[-m(1-p_m)^{m-1} p_m^{n-1} + (n-1)(1-p_m)^{m} p_m^{n-2} \right] &> 0 \\
n(n-1)(1-p_m)^2 - m(m-1) p_m^2 &> 0
\end{aligned}$$

and the right-hand side of the last inequality has a unique root in $p_m \in (0, 1)$ which coincides with p_m^{**}.

Thus, consider different values of B_i's which could satisfy (19) and (20): if $B_i < \hat{B}_i$ then the mixing condition of i crosses the $p_m + p_n = 1$ line twice; if $B_i > \hat{B}_i$ it does not cross the $p_m + p_n = 1$ line; and if $B_i = \hat{B}_i$ it crosses the $p_m + p_n = 1$ line exactly once in p_m^{**}. $\square$

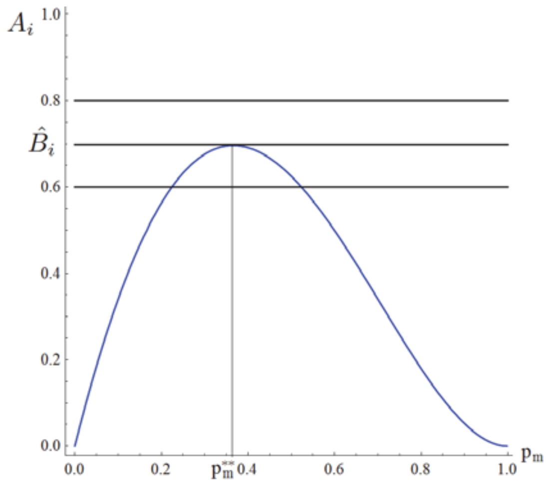

Figure 6: A_i conditional on $p_m + p_n = 1$ as a function of p_m, and its crossings with B_i

Definition 1. *We can express the mixing conditions as a function of p_m as follows: $BR_m : p_m \to p_n$ and $BR_n : p_m \to p_n$ respectively.*

Note that they are both continuous by continuity of the "Mix-Mix" conditions (7) and (8) in both p_m and p_n. Note furthermore that we defined both BR's as functions of p_m into p_n. We do this so we can compare their slopes in the following proposition.

Proposition 3. *For a given (p_m, p_n), $A_m \geq A_n$ if and only if $\frac{\partial BR_n(p_m)}{\partial p_m} \geq \frac{\partial BR_m(p_m)}{\partial p_m}$.*

Proof. We write condition (7) as $A_m(p_m, BR_m(p_m)) = B_m$. Note that we have substituted into p_n $BR_m(p_m)$ and by the implicit function theorem we get

$$\frac{\partial BR_m(p_m)}{\partial p_m} = -\frac{\frac{\partial A_m(p_m,BR_m(p_m))}{\partial p_m}}{\frac{\partial A_m(p_m,BR_m(p_m))}{\partial BR_m(p_m)}}$$

and similarly for individual n

$$\frac{\partial BR_n(p_m)}{\partial p_m} = -\frac{\frac{\partial A_n(p_m,BR_n(p_m))}{\partial p_m}}{\frac{\partial A_n(p_m,BR_n(p_m))}{\partial BR_n(p_m)}}$$

Therefore $\frac{\partial BR_n(p_m)}{\partial p_m} \geq \frac{\partial BR_m(p_m)}{\partial p_m}$ is equivalent to

$$\frac{\frac{\partial A_m(p_m,BR_m(p_m))}{\partial p_m}}{\frac{\partial A_m(p_m,BR_m(p_m))}{\partial BR_m(p_m)}} \geq \frac{\frac{\partial A_n(p_m,BR_n(p_m))}{\partial p_m}}{\frac{\partial A_n(p_m,BR_n(p_m))}{\partial BR_n(p_m)}} \tag{21}$$

and using a similar notation as in the Proof of Lemma 8 for $\tilde{p}_{s+1,m}$ and $\tilde{p}_{s+1,n}$:[26]

$$\tilde{p}_{s+1,m} = \left(\frac{p_m}{1-p_m}\right)^{s+1}$$
$$\tilde{p}_{s+1,n} = \left(\frac{p_n}{1-p_n}\right)^{s+1}$$

[26]For the sake of space we use p_n instead of $BR_n(p_m)$ and $BR_m(p_m)$.

we get that (21) is equivalent to

$$\frac{\sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}\tilde{p}_{s+1,m}\tilde{p}_{s,n}\left(m-\frac{s+1}{p_m}\right)+\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\left(m-\frac{s}{p_m}\right)}{\sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}\tilde{p}_{s+1,m}\tilde{p}_{s,n}\left(\frac{s}{p_n}-(n-1)\right)+\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\left(\frac{s}{p_n}-(n-1)\right)}\geq$$
$$\frac{\sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}\tilde{p}_{s,m}\tilde{p}_{s+1,n}\left(m-1-\frac{s}{p_m}\right)+\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\left(m-1-\frac{s}{p_m}\right)}{\sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}\tilde{p}_{s,m}\tilde{p}_{s+1,n}\left(\frac{s+1}{p_n}-n\right)+\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\left(\frac{s}{p_n}-n\right)} \tag{22}$$

It remains to be mathematically proven the fact that condition (22) is equivalent to $A_m \geq A_n$, but is confirmed by *Mathematica*. The code is available upon request. □

8 Appendix C

This appendix will deal with the case of $m = n$. By the logic of QSNE it is natural to assume that in equilibrium $p_m^* = p_n^* = p^*$. Therefore we will have that $A_m = A_n$.

First note that for $B_i > 1$, $B_j > 1$ Lemma 1 holds regardless of m and n. The rest of the cases will be analyzed using a series of lemmas. Table 1 provides a summary of all the cases and corresponding lemmas (notice that it is symmetric).

	$B_i < 1$	$B_i = 1$	$B_i > 1$
$B_j < 1$	Lemma 12		
$B_j = 1$	Lemma 11	Lemma 9	
$B_j > 1$	Lemma 10		Lemma 1

Table 1: Summary of Cases and Lemmas when $m = n$

Lemma 11. *For $B_i = 1$ and $B_j = 1$, there are only two equilibria: the "Pure-Pure" $(0, 0)$ and $(1, 1)$ for all $i, j \in \{m, n\}$ and $i \neq j$.*

Proof. When $B_i = 1$ then the only way for an individual i to vote with positive probability is when $A_i = 1$ as well, meaning that individual i is pivotal for sure.Thus he must know with certainty the number of individuals i and j that will vote. But this can only happen if i and j individuals are using pure strategies. Let $p_i = 1$ and $p_j = 0$. In this case, an individual i has an incentive to deviate and not vote, so $(1, 0)$ or $(0, 1)$ cannot be equilibria. However, if $p_i = 1$ and $p_j = 1$ then no-one has an incentive to deviate because since $A_i = B_i(= 1)$ an i individual would be indifferent between voting or not, making $(1, 1)$ an equilibrium. For the same reason (voters being indifferent) we can also sustain $p_i = 0$ and $p_j = 0$. Since no-one is voting, everybody is pivotal and $A_i = B_i(= 1)$ still holds. □

Lemma 12. *For $B_i \leq 1$ and $B_j > 1$, the unique equilibrium is $p_i^* = 0$ and $p_j^* = 0$ if the equality ($B_i = 1$) holds and $p_i^* = 1 - B_i^{\frac{1}{i-1}}$, $p_j^* = 0$ when the equality does not hold, for all $i, j \in \{m, n\}$ and $i \neq j$.*

Proof. By Lemma 1 $p_j^* = 0$. If the equality ($B_i = 1$) holds, then the only case for i to vote with positive probability is if $A_i = B_i = 1$. However since no-one from j is voting, the only way to have $A_i = 1$ is by not voting at all. Therefore the unique equilibrium is $p_i^* = 0$ and $p_j^* = 0$. If the equality does not hold, the only case for i to vote with positive probability is if $A_i \geq B_i$. Then $p_i = 0$ cannot be sustained in equilibrium because in this case $A_i = 1$, and any i voter has an incentive to go and vote. Furthermore, $p_i = 1$ cannot support an equilibrium because any i individual would have an incentive to deviate and not vote. Therefore the i individual must be mixing. Plugging $p_j = 0$ into the expression for A_i, we get: $A_i = (1 - p_i)^{i-1}$, so the mixing condition becomes: $B_i = (1 - p_i)^{i-1}$. Solving for p_i we get the equilibrium probability of voting for i: $p_i^* = 1 - B_i^{\frac{1}{i-1}}$.

□

Lemma 13. *For $B_i < 1$ and $B_j = 1$, $p_i^* = 1$ and $p_j^* = 1$, for all $i, j \in \{m, n\}$ and $i \neq j$.*

Proof. Since $B_j = 1$, an individual j votes only if $A_j = 1$, which cannot happen under mixed strategies. Then $p_i^* = 0$, $p_j^* = 0$ cannot be an equilibrium because the i voters have incentive to vote since in this case $A_i = 1 > B_i$ due to the fact that all i voters are pivotal. However, $p_i^* = 1$ and $p_j^* = 0$ is not an equilibrium either because then an i individual would prefer to not vote, and by the same logic, $p_i^* = 0$ and $p_j^* = 1$ is not an equilibrium either. Then the only remaining "Pure-Pure" case is: $p_i^* = 1$ and $p_j^* = 1$. This is an equilibrium because under this we have: $A_i > B_i$ and $A_j = B_j = 1$. □

Lemma 14. *Without loss of generality let $B_i \leq B_j < 1$ for all $i, j \in \{m, n\}$ and $i \neq j$. Then there exists only one "Pure-Pure" QSNE $p_i^* = 1$, $p_j^* = 1$, and at most two "Mix-Mix" equilibria.*

Proof. For the first part, notice that neither $p_i^* = 0$, $p_j^* = 1$ nor $p_i^* = 1$, $p_j^* = 0$ can be equilibria, since members of the voting group have an incentive to not vote. For both of the other two candidate "Pure-Pure" equilibria ($p_i^* = 0$, $p_j^* = 0$ and $p_i^* = 1$, $p_j^* = 1$) we have $A_j = A_i = 1 > B_j \geq B_i$. This means that $p_i^* = 0$, $p_j^* = 0$ cannot be an equilibrium because any individual would have an incentive to deviate and vote, and that $p_i^* = 1$, $p_j^* = 1$ is in fact an equilibrium because nobody has an incentive to deviate. For the second part, since we need $A_i = B_i$ and $A_j = B_j$, the result follows from the analysis in Appendix B.

□

References

[1] Borgers, T. 2004. "Costly Voting." *American Economic Review*, 94(1): 57-66.

[2] Campbell, C. M. 1999. "Large Electorates and Decisive Minorities", *Journal of Political Economy*, Vol. 107, No. 6, pp. 1199-1217

[3] Feddersen, T., and Sandroni, A. 2006. "A Theory of Participation in Elections", *American Economic Review*, Vol. 96, No. 4 (Sep., 2006), pp. 1271-1282

[4] Palfrey, T. R., and Rosenthal, H., 1983. "A strategic calculus of voting", *Public Choice*, vol. 41(1), 7-53.

[5] Palfrey, T. R., and Rosenthal, H., 1985. "Voter Participation and Strategic Uncertainty", *The American Political Science Review*, Vol. 79, No. 1 (Mar., 1985), pp. 62-78

[6] Taylor, C. R., and Yildirim, H., 2010. "A unified analysis of rational voting with private values and group-specific costs", *Games and Economic Behavior*, Volume 70, Issue 2, November 2010, Pages 457–471

[7] Mas-Colell, A., M. Whinston and J. R. Green. Microeconomic Theory. New York: Oxford University Press, 1995.